PIBS- 2891
FILE COPY
APRIL 5/1994

ENVIRONMENTAL RESEARCH

RESEARCH AND TECHNOLOGY BRANCH

ENVIRONMENTAL VALUATION RESEARCH

FINAL REPORT Project No. 428G

1 June 1988 - September 1993

Report prepared by:

Ontario
Environment
Environnement

ISBN 0-7778-2368-3

ENVIRONMENTAL VALUATION RESEARCH

FINAL REPORT Project No. 428G

1 June 1988 - September 1993

Report prepared by:

Jack L. Knetsch
Professor of Economics, and Professor of Resource Management
Principal Investigator

School of Resource and Environmental Management
Simon Fraser University
Burnaby, BC V5A 1S6

APRIL 1994

Cette publication technique n'est disponible qu'en anglais.
Copyright: Queen's Printer for Ontario, 1994

This publication may be reproduced for non-commercial purposes with appropriate attribution.

PIBS 2891

This final report covers research on environmental valuation issues carried out at Simon Fraser University (SFU) with the support of the Ontario Ministry of the Environment (MOE) over the five year duration (with extensions) of the project, 1 June 1988 to 30 September 1993. Individual studies, which the MOE support allowed to be carried out in greater numbers and at a far more extensive scale, are parts of a continuing program of research at SFU on estimation of non-pecuniary values, gain and loss valuation disparity issues, fairness and equity, remedies, and related policy topics. The research program includes an emphasis on implications of the findings and extensive dissemination and discussion of results through publications, conference and seminar presentations and formal and informal courses and workshops.

The timing of the research carried out in this project and the presentation of the findings, has been particularly fortunate. After languishing in largely the domains of scholarly and not so scholarly papers and professional journals, assessments of environmental and other non-pecuniary values are now receiving increased attention and are being taken seriously in making actual choices between resource uses, in designing policies, in setting resource prices, in determining damage awards, and in guiding judicial decisions. The demands for valuation numbers have never been greater.

However, with this increased attention and newly found and rapidly growing enthusiasm has come the need for careful and credible appraisals of important valuation issues. Our studies have centred on these areas, and with little doubt, the findings have had a major impact.

The results of our studies have both significantly advanced understanding of valuation issues, and raised serious questions about the validity of many valuations and the usefulness of the most prominent, and widely promoted, new assessment practices. While many of our results and the conclusions that follow from them have produced some controversy and have not been particularly welcomed in some quarters, they are widely cited and discussed and remain credible and influential.

Although many topics were covered in the research, they can usefully be discussed in this report under two major topics: (1) the disparity between the valuation of gains and losses, and (2) contingent valuation methods. In both areas the studies and discussions included both primary findings and examinations of the policy and planning implications of the results. Dissemination and discussion of the results of the research were a major part of the project and are discussed in a concluding section.

Many of the findings and discussion have been detailed in a series of published and unpublished papers which are included in this report as Attachments. Most of the details of the empirical results and discussions are left to them, rather than repeated in this summary report.

The Disparity Between the Valuation of Gains and Losses

A major issue in environmental valuation, and focus of much of the research, is the recognition of people's common attachment of greater weight to losses than to otherwise commensurate gains. This main finding runs counter to conventional economic dictates and much of current environmental valuation practice. It is, however, a subject which in recent years seems finally to have attracted a general and sympathetic interest among economists -- though, curiously, not yet among most people dealing with resource and environmental valuation questions.

The main issue is that there are two measures of economic value -- one for gains and another for losses -- and while we have long assumed that they would yield essentially equivalent assessments, the empirical evidence suggests otherwise. There is no dispute that economic value of both gains and losses are measured by what people are willing to sacrifice. Gains are valued in terms of what people are willing to pay (WTP) to obtain them, and losses are appropriately valued in terms of what people would demand to accept them (their willingness-to-accept, or WTA). The usual definitions are some variant of: "the economic value of something is how much someone is willing to pay for it or, if he [or she] has it already how much money he [or she] demands to part with it" (Posner, 1986, p.11).

Although income, or wealth, effects or limits on the ability to pay may cause the two measures to differ, these are normally not a factor of much practical importance and it is agreed can safely be ignored in most applications. The common view is that, "for many goods, services, and amenities that command a modest fraction of the consumer's budget, the differences between [the] ... measures are trivial" (Randall, 1987, p.244). Consequently, the usual advice is that "as a practical matter it usually does not make much difference which of these two approaches ... is adopted" (Rhoads, 1985, p.125).

In spite of the agreement that losses are appropriately valued by the WTA measure, resource and environmental losses are almost invariably valued in practice by the WTP measure. This has been justified on grounds of both convenience -- and the difficulty of estimating WTA -- and these assurances of equivalence between the measures. However, the results of essentially all of empirical tests indicate that people's preferences and actual choices are not at all consistent with the equivalence assertion. Alas, people value losses far more than gains, and the common practice of assessing losses by the WTP measure will result in assessments that greatly understate them.

Evidence of a Disparity

Large differences between the two measures were first noted in survey studies of people's valuations of various losses of environmental assets or the degradation of environmental quality. For example, individuals in a sample of duck hunters said they would be willing to pay an average of $247 to save a marsh area used by ducks, but would demand an average of $1044 to accept the identical loss (Hammack and Brown, 1974). There are serious reservations about whether or not these earlier studies really estimated economic values, as such, but they did alert some of us to the real possibility that people did indeed values losses differently than gains (Gordon and Knetsch, 1979).

This led to later studies based on real exchange experiments -- the first conducted in the early 1980s (Knetsch and Sinden, 1984) -- which have provided more stringent tests. The results, however, have been essentially the same. Even when exchanges of real goods and actual cash payments motivated the evaluations, the compensation demanded to give up an entitlement far exceeds the comparable payment measure of values.

An indication of the general finding is illustrated by the results of a simple experiment, carried out as one of the present research studies, involving people's actual choices between a coffee mug and a large, four hundred gram, chocolate bar.[1] Each member of one group of 76 participants was given a mug and told that it was theirs to keep and take with them. They were then given a chance to give up their mug to obtain (that is, gain) a chocolate bar. A second group of 87 individuals were given the opposite choice -- to give up chocolate bars to obtain coffee mugs. At the end of the exercise, participants from both groups could end up with either a mug or a chocolate bar, in accord with their own individual preferences.

The traditional assertion of equivalence of gain (WTP) and loss (WTA) measures yields the strong prediction that, given the absence of any significant transaction costs, about the same proportion of participants in each of the groups will prefer mugs to chocolate bars, and will therefore choose accordingly. The actual results were in sharp contrast to this prediction: 89 per cent indicated a preference for mugs when initially given a mug, and only 10 per cent revealed a similar preference for mugs when they had to give up (lose) a chocolate bar in order to obtain (gain) one. The relative value of mugs and chocolate bars varied greatly and depended on whether the evaluation was made in terms of a gain or a loss. The mug was valued more when it had to be given up to obtain a chocolate bar; and was valued less when it was obtained in exchange for giving up a chocolate bar. Gains and losses were valued very differently.

---

[1] Details of this particular study are described in Knetsch (1989), included as Attachment A.

Much of the early project research was focused on testing the persistence of the valuation disparity under widely varied conditions. While there were many reports of differences, questions were also raised about their significance given the possibilities of being limited to largely artificial settings and one-time transactions. An extensive series of experiments and other studies were conducted to provide repeated tests of the impacts of transaction costs, the nature of the entitlements, repetition and learning, and income or wealth effects. The results were basically unchanged -- evaluation disparities persisted over all of the tests, including the critical ones of repetition.

These findings were presented to a wide array of conference and seminar participants, in order to review and anticipate the professional scrutiny that they were sure to encounter. After these discussions, a paper summarizing the results was eventually accepted, after further intensive peer review, and published in the Journal of Political Economy (Kahneman, Knetsch and Thaler, 1990; and Attachment B). This is not only one of the two or three leading professional journals in the field, but is one generally viewed as being more likely to accept reports of more traditional economic research findings. We regard the results reported in this paper, and its publication in the JPE, as major accomplishments of the project.

The real exchange experimental tests, and those of others, have not only affirmed the persistence of evaluation disparities over repeated valuations, but they have also fairly well precluded the possibility that the differences in the valuations of gains and losses might be attributable to transaction costs or strategic behaviour on the part of participants. These tests, however, have used between subject designs therefore limiting the results to apply to a representative individual. While not a major concern, even this worry has been tested and similar differences in the valuations of gains and losses by the same individuals have also been demonstrated by Kachelmeier and Shehata, (1992). In all, the results of the empirical tests indicate that the differences are pervasive and large, and likely not an artifact of any particular study.

People have also been shown to exhibit similar disparities in their valuations of gains and losses in their actual behaviour in making ordinary real choices -- findings which we, in part, summarized (Kahneman, Knetsch and Thaler, 1991; Attachment C). Frey and Pommerehne, for example, have noted the large asymmetry between public support for saving national art treasures and for acquiring ones of equal or greater artistic merit (1987). The reluctance to sell at a loss, attributable to the valuation disparity, has been found evident in the higher volumes of house sales when prices are rising and the smaller volume of sales of securities that have declined in price relative to those for which prices have increased (Shefrin and Statman, 1985).

An interesting illustration of the valuation disparity appeared in the choices of automobile insurance policy options made by people in the U.S. states of New Jersey and Pennsylvania. In both states auto owners could choose between a cheaper policy which limited rights to further recovery of damages, or a more expensive one which allowed that possibility. Although the difference in premiums was substantial and little effort was involved in making a change, overwhelming majorities in both states chose the default option rather than give it up for the alternative, even though the limited rights policy was the default in New Jersey and the more expensive was the default in Pennsylvania (Johnson, et.al., forthcoming).

Although findings of valuation disparities are now quite commonplace, the early studies reporting large valuation disparities tended to generate a good bit of skepticism, and certainly little inclination to take this evidence into account in assessing resource and environmental values. While quite appropriate questions were raised about methods and other aspects of the evidence, the major reason for the present reluctance to take the results seriously appears to be that they were inconsistent with long accepted economic theory -- a slight perversion of the usual course of scientific inference.[2] However, the basic findings have now been replicated by many investigations using widely varied research designs, and the results are now commonly reported in what are acknowledged to be the highest ranking journals in the field. The disparity between measures of gains and losses is clearly no longer a curiosity. Instead, the pervasive difference appears, contrary to traditional views, to simply be a common characteristic of people's preferences. This is, after all, largely an empirical matter and it now turns out that a characteristic value function is better described as one having a kink at the reference point rather than one that is increasing at a smoothly decreasing rate (Kahneman and Tversky, 1979). The larger question now seems to be the extent to which these findings should be taken into account in valuation studies, as well as in other economic analyses, and in the design of public policies and planning of private actions.

---

[2] This is not inconsistent with the early experience with economics which Herbert Simon, winner of the Nobel Prize for Economics in 1978, recalls in his autobiography: "And since my exposure to the economics profession was still rather minimal, I had not acquired the habit, so common in that profession, of ignoring the real world when it contradicts the theory" (1991, p.371).

Implications of the Valuation Disparity

The differing weights attached to gains and losses have a number of important implications. These include ones for many applications of economic analyses as well as ones more directly involved with measurement of environmental and other resource values.

Non-Reversibility of Indifference Curves

A basic assumption of standard economic principles is that indifference curves are reversible, that is, the rate at which people will sacrifice one good for another is presumed to be independent of their initial entitlements and the direction of exchange offer. This assertion is, however, violated to the extent that giving up a good has a greater impact on an individual's welfare than gaining the same entitlement. Such non-reversibility has been demonstrated for several routinely traded goods, in studies which were part of the present project (Knetsch, 1989 and 1992; and included as Attachment A and Attachment D). These findings of a non-reciprocal relationship, which have been shown for common goods and goods and money, might now, on the basis of these results, be expected to be common to many other trade-offs as well.

Preference Order Axioms

The disparity observations also call into question several axioms of preference orderings that are presumed to provide major support for much of contemporary demand and valuation theory and a good bit of economic analysis generally. Preferences to keep a good A rather than give it up for another good B, but also to keep B rather than exchange it for A if given the opposite choice, is largely inconsistent with dominance, independence, completeness, and transitivity criteria. All of these violations were demonstrated in a real exchange experiment in one of the research studies (Knetsch, 1993a; and Attachment E).

Coase Theorem

The proposition which has become known as the Coase Theorem, suggests that the assignment of initial entitlements to one or the other party will not matter if the two parties are allowed to engage in costless voluntary exchanges -- the final arrangements of entitlement and liability will be efficient and independent of the original allocations. This surprising result, published in what has become the most cited paper in the economics literature of the past 50 years, has become a prominent stricture of economics and increasingly is used in the analysis of legal rules and design of public policy reforms.

However, the dependence of valuation on the reference position also undermines the primary prediction of the Coase

Theorem. Even in the absence of transaction costs, voluntary exchanges will not necessarily assure that final arrangements of entitlements will be efficient and the same regardless of initial assignments -- as a receipt foregone is often not treated the same as a payment of the same amount. This point was empirically demonstrated in Kahneman, Knetsch and Thaler (1990; Attachment B).

Fairness and Acceptance

We found, in earlier studies, that the disparity between gains and losses also appears to influence people's judgments of what they regard as being either fair and acceptable or unfair and unacceptable economic behaviour. If an action is seen to impose a loss on one party for the benefit of another, this will nearly always be widely seen to be unfair -- quite apart from whatever economic justification might be thought to exist. For example, raising prices in response to sudden shifts in demand is seen to benefit the seller, who has not incurred any cost increase, at the expense of the buyer and is therefore judged to be unfair. Similarly, cutting wages when unemployment increases is thought unfair by the majority of people because the employer benefits in direct proportion to the workers' loss and the impact of the loss is taken to outweigh that of the gain. Reducing a customary bonus payment to workers is apparently viewed as a less aversive foregoing of a gain and is therefore far more acceptable than an equivalent reduction in wages, which is commonly seen as imposing a loss on workers (Kahneman, Knetsch, and Thaler, 1986a and 1986b).

Judicial Choice

Judicial choices also appear to reflect much more sensitivity to the disparity between gain and loss valuations, and the standards of fairness and acceptability that stem from them, than most conventional economic analyses would suggest. A study of judicial choices in common law jurisdictions revealed such a pattern. For example, recovery of expenses resulting from wrongful actions, which is taken as a direct cost or loss, is far more likely than compensation for forgone profits, which appears to be viewed as a less important foregone gain. For much the same reason, court decisions show a consistent bias of reluctance to shift entitlements from current holders (Cohen and Knetsch, forthcoming; and Attachment F).

Negotiation and Conflict Resolution

The disparate views of losses and foregone gains may also influence the acceptability of alternative negotiation and conflict resolution proposals. Early results suggest that judgments of acceptance are far more sensitive to direct costs than they are to opportunities foregone (Borges and Knetsch, 1993; and Attachment G). This and other behavioural findings offer encouraging promise of greater understanding of which characteristics make proposals for resolving conflicts more acceptable, and which discourage settlement. Offers might then be better designed to be more sensitive to people's real, and not just assumed, concerns.

Mitigation and Compensation Remedies

People's reactions to proposed remedies for harms imposed on them provide another interesting, and potentially important, implication of the valuation disparity. The common prescription draws on the usual economic critique that presumes that people should favour compensation over efforts to mitigate the injury, because a money award can be used by people to buy whatever they value most whereas mitigation restricts their remedy to the one of only reducing the specific adverse impacts of the action. This turns on the well-known assertion that a monetary award will yield greater welfare gains than an equal sum in the form of a particular good -- the usual textbook demonstration allegedly shows the superiority of money over housing allowances or food vouchers.

However, the behavioual findings suggest that because mitigation measures reduce losses they will therefore be valued more highly, and because compensation will likely be viewed as a gain it will be discounted. In terms of the more realistic kinked value, or welfare, function, mitigation efforts increase welfare assessed by the steep portion of the curve in the domain of losses, and compensation results in much more modest increases in welfare indicated by the relatively flat response in the domain of gains. Further, the findings of what people regard as fair and unfair, suggest that tying a relief action to the injury, as with the acceptance of passing on costs, greatly increases the acceptability of a remedy. The data from one of the current studies, and discussed in Knetsch (1990; and Attachment B) are consistent with these findings, showing that people do indeed often prefer mitigation over compensation.

The findings of pervasive valuation disparities suggest that more attention to the design of remedies may well be in order. Common directives to pay off parties to "make them whole", and to eliminate opposition to various development projects and activities, for example, may be expensive and ineffective responses. Remedies better tailored to the perceived dimensions of particular losses are likely to be more effective and efficient than approaches formulated on the basis of more traditional behavioural assumptions. This might take the form of determining what attributes, or dimensions, of gains and losses people view as being affected by an actual or anticipated change, and designing remedies to more closely offset the adverse impact on these attributes.

Sustainable Development

The restrictions on the acceptance of substitutes, implied by the gain/loss valuation findings, raises a possible related implication for policies designed to achieve sustainable development. One prevalent notion of sustainability turns importantly on the acceptability of substituting one form of natural (or even human) capital for another, thereby maintaining a desired level of productivity -- a version sometimes referred to as "soft sustainability". Thus if forests are harvested in one area, total productivity might be maintained by planting an alternative species of trees in another region or by enhancing fish or wildlife habitat. However, the preliminary data suggest that people's willingness to substitute one resource gain for another resource loss, may be more constrained than usually presumed. This possible restraint on acceptable substitutes -- due to the disparity between gain and loss valuations -- may influence the design of sustainability policies.

Dedicated, or "Ear-Marking" Funds

In much the same way that the gain and loss disparity limits acceptance of substitutes, there appears to be a public finance lesson in this for the use of funds collected from user fees or other special purpose levies. Just as people often prefer mitigation measures to monetary compensation for an injury, they also want such monies used for a purpose related to the levy -- the widely denounced "earmarking" of funds -- rather than put into a general or consolidated revenue account. And for much the same reason: the funded activity reduces the adverse feeling of the loss if it is seen to be related to the purpose for which the money was taken. The usual public finance dictum that such funds should be put into the central accounts and used for any purpose most valued by the community, might be in need of some amendment to better reflect the real valuations of the community in these matters.

Environmental Valuations

In addition to the other implications discussed, the valuation differences have, of course, very direct consequences for valuation practice. The usual procedure of using WTP measures to assess losses will result in systematic, and on current evidence, very serious underestimates of their value. Consequently, activities with negative impacts will be unduly encouraged as the true adverse impacts will be underestimated; compensation and damage awards will be under assessed; inappropriately lax standards of protection against injuries will be set as the added costs of further harms will be heavily biased toward under assessment; inadequate mitigation measures will be undertaken as the benefits of preventing further losses will be incorrectly measured; choices of legal entitlements will be biased because comparisons between the efficiency of alternative allocations will be based on incorrect measures; too few resources will be devoted to avoiding injuries; and in general full accounting and appropriate pricing of non-pecuniary inputs will be frustrated (Knetsch, 1990 and 1993; Attachment H and Attachment I).

An argument can be made that a WTP measure of a loss value may be adequate in cases where a critical threshold value has been established. Such might be the case, for example, if the value of a development is $X net of a particular resource cost and a proposed change would, consequently, be uneconomic if that cost exceeded the $X. If, in such instances, a WTP value was sufficient to demonstrate that the cost exceeded the value of the alternative, it would then provide useful information. In such cases this might be an appropriate valuation strategy.

In general, however, the use of WTP measures is to pose the wrong question. Even in cases for which a critical threshold has been established, no conclusion is possible if the WTP value fails to provide the needed demonstration and an appropriate WTA measure is not provided. And in many cases -- in assessing damages, for example -- a threshold is simply not available. Nor is it a very defensible position to claim that we should use WTP to measure losses because it provides a conservative value. To provide efficient allocations, proper restitution, justifiable deterrence, and correct incentives, we want estimates that are more likely to be correct rather than ones which we know to be wrong.

The Reference Point

The disparity between the measures has the further implication that the choice of measure has a very substantial impact on assessments, and may in many cases determine the outcome. Current practice in making this critical selection is one largely of confusion and indifference.

A recent study has resulted in a tentative set of criteria which could lead to more realistic choices that would more appropriately assess welfare changes (Knetsch, 1993b; and Attachment J).

Contingent Valuation Methods

In addition to commonly seeking answers to the wrong question in valuation of losses, which is suggested by the results of the valuation disparity studies, there is also the further worry about methods of establishing monetary assessments of losses of non-pecuniary values. While techniques such as the travel cost method, hedonic pricing, and replacement cost, are frequently used and produce useful value estimates, they are also widely acknowledged to have limited applicability over the broad range of resource and environmental issues for which valuations are a consideration.

Contingent valuation methods (or CVM), based on survey respondents's statements of value, are alleged to be free of such restrictions and are therefore widely used for estimating resource and environmental losses. It is the use of CVM that gives rise to newspaper accounts headed by statements such as:

> "Australians Want to Save Kakadu for $53 a Year"
>
> "State Residents... Indicated They Would Pay Up to $10 per Bird for Protecting Them From Future Spills"
>
> "Studies Have Indicated that the Cost of the Environmental Damage from the [Exxon Valdez] Spill is Between $3 Billion and $5 Billion"

Further, estimates derived by asking people how much they value whatever is at issue are taken to produce measurements that are fully comparable to economic values derived from competitive market exchanges and thus to allow direct comparisons of the economic values of alternative resource uses. The now common assertion, or at least hope, of those using CVM estimates is that:

> "[CVM] is potentially capable of directly measuring a broad range of economic benefits for a wide range goods, including those not yet supplied, in a manner fully consistent with economic theory." (Mitchell and Carson, 1989, p.295).

Thus, in cases of allocating a block of old growth forest to either timber sales or wilderness on grounds of maximizing economic value, the market value of timber is assumed to be directly compared to the CVM estimate of the wilderness value.

The now immense popularity of CVM to estimate non-pecuniary resource and environmental values is a fairly recent phenomenon. The first application of the method is generally attributed to Robert Davis (1963). While he demonstrated most of the attributes which present advocates proclaim, little serious use was made of the method until recent U.S. legislation made people responsible for environmental degradation -- through oil spills and the like -- financially liable for the value of the environmental loss. The sudden demand for a number on which to base damage awards turned attention to CVM, the method which promised delivery of such numbers.

Given the legitimacy stemming from Court approvals, at least in the U.S., and from a profusion of favourable reports in professional journals, the method is now being used for all manner of purposes in countries throughout the world, Canada certainly included. And its uses -- and fees for carrying them out -- continue to grow at rates seldom seen among adaptations of academic scribblings.

There are now, however, pressing reasons for some sobering reassessment of this enthusiasm. The results of our recent tests, and increasingly those of others, suggest that the CVM results are numbers, but these are not likely to be measures of economic values. Further, the methods also appear to be subject to serious manipulation -- generally regarded as not being a good feature of a method on which damage awards and allocation decisions are to be based.

Anchoring

One problem is that CVM estimates can be very susceptible to anchoring biases, or what is known in the CVM literature as starting point bias. The difficulty can be illustrated by the results of one of our early studies.

The sequential questioning format was, at the time, the variant of the contingent valuation method generally most favoured by users of CVM and commentators writing on the subject. In this procedure respondents are asked a "Yes or No" question of whether or not they would be willing to pay a specific sum for the purpose described to them. If they say "Yes", then the question is repeated using a higher sum; if they say "No" the second question contains a smaller amount. This is repeated until a final WTP number is produced.

We used this procedure, but in our case we split the sample of Toronto residents into different sub-samples and varied the sum they were asked to pay in the first question with respondents in all groups then asked if they would be willing to pay $50 in a second question. That is, one group, for example, was first asked if they would be willing to pay $200 to preserve fish populations in the Haliburton-Muskoka area of Ontario; those saying "No" were then asked if they would be willing to pay $50 for the purpose. A second group was asked if they would pay $25, and those saying "Yes" were asked if they would pay $50. Nothing but the sums was varied in the questions. In this way we were able to estimate the proportions of respondents in each group who said they would be willing to pay $50. If the method is to produce credible assessments, then presumably we should find little difference in these proportions regardless of the amount asked in the first question.

The actual responses indicated anything but a constant rate. Of the respondents first asked if they would pay $200, 63 per cent indicated a WTP of $50. Of those first asked if they would pay $25, only 18 per cent said they would pay $50! People were quite obviously anchored on the initial sum, and the resulting "valuations" were varied accordingly. We have summarized these results in the context of later survey format proposals in Kahneman and Knetsch (1993; and Attachment K). We did similar studies, using other wording, goods, etc., with about the same results.

Partly as a result of this type of finding of large impacts of anchoring, many CVM users have now changed their position to favour a "referendum" or closed ended format. In this procedure, people are asked a WTP question for a single sum; but this sum is varied among respondents. An average WTP can be calculated by (essentially) observing the proportions of respondents who are willing to pay the varied amounts.

Recent tests, however, are pointing to problems with this method as well. In comparing open-ended responses of how much people would pay for a good, the referendum responses show much less variation across the different amounts. That is, the proportions of people willing to pay the amount posed in the question does not differ very much for the varied sums. This results in a much flatter response function, usually indicating much higher average WTP than for other response formats. For example, Kealy and Turner (1993) recently reported the results of a comparison between referendum and open-ended formats and found that the former procedure produced estimates about twice as large as the other. (They may inadvertently have admitted more futility of the whole exercise in suggesting at the end of their report that in such situations "where the estimate is sensitive to question method, one should report both estimates" -- a fairly frank admission that analysts really do not have much idea about the usefulness of CVM results.) An earlier similar finding was reported by Desvouges, et.al. (1992). And our further tests suggest that the results from the referendum, or any other format, may well be even less credible.

Embedding

While there are other limitations and worries that make CVM estimates suspect in many cases, the problem of what has become known as embedding is turning out to be perhaps the most serious. This is the case when the same good is assigned a lower value if WTP for it is inferred from WTP for a more inclusive good rather than if the particular good is evaluated on its own.

We first noted this problem in an early series of surveys when we asked a random sample of Toronto residents if they would be willing to pay $50 to maintain fish populations in particular areas. One sub-sample was asked if they would pay to maintain fish in a small area of Ontario; a second sub-sample was asked about paying for fish in a larger area of the Province, that included the area asked of the first group; and a third sub-sample was asked for their willingness to pay for fish in the entire Province. The, then, surprising result was that the proportions of people willing to pay $50 to save the fish did not vary much with the area mentioned in the question, in spite of the vastly different numbers of lakes and streams, and consequently fish, that would be protected.

Although others had noted related phenomenon, which they variously termed a part-whole effect, symbolic effect, or disaggregation effect (see, e.g., Cummings, et.al., 1986; Mitchell and Carson, 1989), and Kahneman reported our Ontario results at an important "state of the art" CVM conference in 1984, these findings had little or no impact on either evaluation practice or discussions.

We, however, continued to look into the issue, mainly because we thought it interesting. One idea we tested was that the lack of difference in values of larger or smaller numbers of fish, might be due to feelings that saving fish was important for symbolic purposes and that the quantity mattered little. To do this we asked different sub-samples about paying $50 for two different and unrelated goods, and for a package of the two. For example, we asked one group if they would pay $50 for a weekend holiday at an Ontario resort, a second group $50 for coupons for groceries, and a third $50 for both the resort holiday and grocery coupons. The results, in terms of the proportions saying they would pay $50, were:

Weekend at Ontario resort	59%
Grocery coupons	60%
Resort + groceries	59%

Another similar test involved a weekend in Montreal, and gift certificates for a linen and towel shop; goods that did not seem to be at least obvious substitutes. These results were even more curious:

Linen shop certificates	24%
Weekend in Montreal	61%
Linens + Montreal weekend	55%

The combination of the two goods was valued if anything, a bit less than one of the goods alone.

We then did another test using two public environmental goods which did not seem to be substitutes for each other. In this case people in each of two subsamples of Vancouver residents were asked an open-ended question of how much they would pay for one of the public goods, and a third sub-sample was asked their willingness to pay for a combination of both goods: maintaining fish populations in the interior regions of British Columbia, and protection of historic buildings throughout the Province. The results of the average WTP, again, did not conform to normal valuation expectations:

Fish populations	$52.35
Historic buildings	55.57
Fish + historic buildings	51.10

Clearly, both goods are valued positively, yet people seemed unwilling to pay more for both goods than they would pay for either one of them alone (in fact, slightly less).

Later, motivated in large part by our belated realization of how seriously CVM was being taken, we designed and conducted a more thorough test for embedding. We divided a sample of Vancouver random household respondents into three groups. The first were asked how much they would pay for a group of public goods, then how much of that sum for a portion of these goods, and finally how much of that for the least inclusive of these. Individuals in the second group were first asked how much they would pay for the portion of the list of public goods that made up those in the second question put to the first group of respondents; they were then asked what portion of that sum they would pay for, again, the same least inclusive public good. Respondents in the third group were simply asked their WTP for the identical least inclusive public good -- availability of rescue equipment and trained personnel. The means of the responses showed major embedding:

	Group 1	Group 2	Group 3
Services including disaster	$136		
Disaster including equip/pers.	29	$152	
Equipment/personnel	14	77	$122

We reported these, and some related results, in what turned out to be an important but controversial publication (Kahneman and Knetsch, 1992a; and Attachment L). The paper clearly raised interest in the topic, and was widely distributed and discussed even before its publication. (The report has also prompted some disappointing untoward reactions and behaviour among many in the resource and environmental valuation community.)

The most pertinent result of our test is that the same good, for example equipment and trained personnel, is valued very differently depending on its being valued alone or initially as part of a more inclusive set of goods -- a difference in this case of ranging from $14 to $122. We have called this "regular embedding" to distinguish it from another common result of "perfect embedding", which refers to the common tendency of people to give similar WTP responses to more or less inclusive goods -- evident in this case by the lack of a statistically significant difference between the $136, $152, and $122 responses (Kahneman and Knetsch, 1992b; Attachment M).[3]

These basic results have now been replicated in further studies. Perhaps the most notable (or notorious in some quarters) of these is a series of tests carried out by a group of prominent economists, econometricians and psychologists, supported by the EXXON Corporation and initially reported at a conference in 1992 (Cambridge Economics, 1992).

The demonstrated common presence of embedding in CVM responses seriously undermines the credibility of taking these responses as measures of economic value. The contrast between indicators of economic values and responses to a CVM survey in which embedding is present can be illustrated by a simple hypothetical case of a well-informed shopper desiring two common goods, milk and muffins.

---

[3] This paper is a Reply, prepared to respond to a lengthy invited Comment which was published after our paper. This was published, as cited. A second much more hostile, and we believe mistaken, Comment was subsequently published as well. We prepared a Reply (Kahneman and Knetsch, 1992c; and Attachment N), as is customary in such cases. However, in a development that seems to be indicative of the present tenor of what has been termed in the popular press as an "academic brawl" by one of the participants, our Reply was surprisingly -- and even astonishingly -- rejected by the editors of the journal!

The shopper can be presumed to be aware that milk is readily available from other convenient sources at a price of $2 per carton, and muffins are priced at $1 each at a variety of similarly accessible shops. Given the ready availability of substitutes, the most that this shopper would be willing to pay -- to sacrifice -- for a litre of milk at any particular site or store would then be $2. This maximum willingness to pay of $2 is the economic value this shopper places on milk from a particular store. It is not necessarily the maximum amount that this individual would pay for milk per se, in the absence of substitute sources, but it is the maximum valuation given the alternatives that exist.

The same is true of muffins. The economic value of a single muffin from a particular shop is $1, given their availability at this price from alternative sources. And given the absence of any discount for buying two, the maximum willingness to pay for two muffins is $2.

An elicitation of the shopper's true economic valuation of these goods would reveal that the most that he or she would be willing to pay for a single muffin is $1, and the maximum willingness to pay for two is $2. Similarly the most that would be paid for two muffins plus a litre of milk would be $4, and $2 of that total would be attributable to milk and the other $2 to the two muffins. This case involves no embedding and (following our demonstration of embedding) answers to alternative valuation questions put to the shopper can be represented as follows:

Willingness to pay for:

Milk + 2 Muffins	$4		
2 Muffins	$2	$2	
1 Muffin	$1	$1	$1

Thus, if the shopper were asked the maximum sum he or she would be willing to pay for a carton of milk plus two muffins, the reply would be $4; and if asked how much of this would be the value of the two muffins, the response would be $2; and $1 would be the stated sum attributable to one muffin. Alternatively, if the same shopper were asked the valuation of only two muffins the honest reply would be $2; with $1 the value allocated to a single muffin. And if the shopper is only asked the value of one muffin, the answer would be $1. The value of two muffins is revealed to be $2, whether asked as a single transaction for two muffins alone or asked as part of a bundle with a carton of milk. And the value of a single muffin is $1 whether considered by itself, as one of two muffins, or as part of a bundle with another muffin and a carton of milk.

The presence of embedding, so strikingly evident in CVM surveys, reveals a very different pattern of "evaluations". Responses typical of the empirical evidence of embedding, might be these:

Willingness to pay for:			
Milk + 2 Muffins	$4		
2 Muffins	$3	$4	
1 Muffin	$1	$2	$4

Thus, the shopper might say that the maximum willingness to pay for the bundle of two muffins and a litre of milk is $4, but that $3 of that is for the two muffins, and of that $1 is for a single muffin. Alternatively, if asked the value of just two muffins, the response might also be $4, with $2 the value of just one of these; and if asked the maximum sum for a single muffin alone the answer might well be $4. The value of two muffins can then be taken as $3 or $4, and the value of a single muffin $1, $2 or $4.

Unlike the elicitation of true economic values that reveals consistent evaluations regardless of the form of the questioning, CVM responses in the presence of embedding show a total lack of consistency. The indications of value -- the CVM responses --differ depending on how the good is combined with or without other goods. And this occurs, as in this illustration, even in the absence of any substitution possibilities between the goods or any problem of satiation.

Some advocates of CVM estimates have acknowledged the pervasiveness of embedding in the responses, but have attempted to defend the validity of the results as being economic values on the grounds that if the question is context specific to the gain or loss at issue then there is no reason to be concerned that this value would change if the context were changed to be part of a more or less inclusive gain or loss. So if the change is two hectares of wildlife habitat then the question should be asked about just two hectare of habitat and not about one or three hectares or about two hectares plus a tennis court. If this is done then the responses are alleged to represent the value of the two hectares of habitat in a form fully comparable to and compatible with values derived from well functioning markets.

However, this assertion appears completely specious. The CVM answer is but an artifact of the question, not an indication of the value of the good at issue. If, for example, the gain or loss at issue[4] is two muffins, a question concerning the value of two muffins would reveal a "value" of $4. But if the replies are to be taken seriously as economic values, and in the absence of satiation or other reasons for diminishing marginal values, then the question could also be asked of the value of a single muffin and the response doubled to determine the value of two muffins. Such a procedure would, given the embedding indicated in the example, yield a value of $8 for the two muffins. Or if the value is derived from the indicated portion of the value of a more inclusive group of goods of two muffins and a litre of milk, the "value" would then be given as $3. And the "value" of two muffins can also be taken as $2 if derived from a doubling of the value of one muffin when this is taken as the portion of the value of a bundle of two muffins and a carton of milk attributable to the one muffin.

---

4. Maintaining the neutrality between a gain or a loss is simply to avoid the issue of losses being valued much more than commensurate gains -- a difference that purveyors of CVM studies inappropriately, and at great social cost, continue to

If the shopper is faced with the prospect of being given either a bottle of juice which he or she values at, say, $4.50 or a muffin then presumably welfare and efficiency is best served by a choice of the juice -- it is worth $0.50 more then indicated value of the muffin. However, if the shopper is instead faced with the choice of receiving either a bottle of juice or two muffins, then (again, in the absence of satiation) the welfare maximizing choice is ambiguous given the CVM results. The gain of two muffins might be worth twice the $4 of one, or $8 -- or $3, or $2 -- and little judgment of the allocative efficiency of giving either the juice or the two muffins is possible.

Rather than economic values, CVM responses seem more likely to be an indication of an attitude or good feeling of moral satisfaction (Kahneman and Ritov, 1993; Kahneman and Knetsch, 1992a). Consequently, given the presence of embedding, CVM results on present evidence appear to offer little or no guide to allocation policies or basis for damage assessments. On present findings the results of CVM studies cannot be compared to economic values derived from market exchanges and allocative decisions made accordingly, nor does it seem fair to subject people to compensation orders or allocations on the basis of these numbers.

---

ignore.

Conclusions

As suggested initially, the results of our research, as well as that of others, are now calling many current resource and environmental valuation procedures into question. While the enthusiasm for valuations has perhaps never been stronger, the evidence suggests that we may indeed be getting misleading answers to the wrong question.

The persistence in continuing to use WTP values to measure losses, instead of the more appropriate WTA, almost certainly will seriously bias assessments and distort and undermine policies based on them. And CVM studies produce numbers -- this guarantee itself may be one of the real disadvantages of such procedures -- but they do not appear to be economic values as supposed.

There seems little reason, therefore, to agree with the Ontario Law Reform Commission (1990), in advocating that damages for environmental harms be assessed using CVM to estimate WTP values:

> "For example, a sample of people may be asked what they would be willing to pay to preserve the grizzlies in Alberta or to keep an oil slick off Cavandish Beach in Prince Edward Island" (p. 47).

The WTP valuation will likely understate the loss, and CVM is unlikely to produce an economic value at all.

While the evidence points to change, most practitioners of environmental valuations, and eager consumers of their findings, remain in a state of denial -- not a particular good stance to take in the midst of growing demands for the kind of rational and reasoned responses that evaluations are presumably intended to guide. These socially desired interests would likely be better served by a bit more even handed regard for empirical findings and for more willingness to explore alternatives to current practices.

One strategy that might well be worth some quite serious exploration is to value what can be valued using procedures agreed to be credible, to back off from valuations that are with current methods not capable of acceptable monetary assessment, and to rely instead on some form of "interim schedule" of compensation awards or allocation guides. Such a schedule would not need to provide detail for all possible resource or environmental losses, but could provide sufficient benchmarks to provide guidance for specific losses as they arise.[5] These benchmarks could be very rough and yet may be widely accepted by people because they would not be intended to be values, but only guides for compensation and allocation decisions. The evidence from such areas as injury awards for workers compensation claims in Canada and in other countries, and awards for pain and suffering from Canadian courts, which are based on equally rough guides that also do not pretend to represent actual monetary

---

5. A hint of such a proposal is contained in the recent report of the U.S. NOAA Panel report, in their suggestion that "the federal government should produce standard damage assessments for a few specific reference oil spills..." (1993, p.4609). The origin of this proposal is not clear as Danial Kahneman presented the results of some of our discussions of this to the Panel at one of their hearings.

values, is encouraging. These guides are widely accepted, and they serve most of the purposes for which damage awards are designed to serve.

Environmental and resource loss schedules might serve equally well. They could be implemented quickly, they would provide deterrence incentives, they would provide predictable outcomes, they would likely provide widely accepted degrees of restitution, and they would apply equally to all and thereby have greater appeal as being fair.

The design of the schedules -- which should be termed interim on grounds that our fairness findings suggest that a permanent designation would be less acceptable -- would probably not be without controversy, but might not be that difficult. While we are very short of good estimates of absolute values, we usually do have wide agreement on relative values. And this may be largely sufficient for this task. There are also grounds for encouragement that newer decision analysis and value structuring techniques -- unlike CVM -- might provide means to reach wide consensus on the needed schedules (see, for example, Gregory, et.al., 1990; Edwards and Winterfeldt, 1987). At this point, this seems worthy of at least a portion of the support and attention being given to present resuscitation efforts on less promising approaches.

One of the main findings that give rise to this questioning of present valuation methods and procedures, is itself one of the explanations for the observed reluctance to take the new evidence more seriously: namely, the greater value people associate with the loss of things, or ideas, relative to the perceived worth of new things or ideas to replace them. However, the evidence also suggests that some people do give up their initial holdings; and to the extent that resistance continues, it may be time to turn to others less burdened by entrenched reference positions.

Dissemination and Discussion of Research Findings

Extensive presentations and discussions of the results of the research became a prominent feature of the project. Given what many considered the unusual, if not somewhat controversial, findings, it was important to discuss these widely to understand and deal with concerns, and to design further studies to test suggested explanations and extensions. And given the potential importance of many of the results, it was also important that they became widely discussed and policy analysts and environmental planners be made aware of them -- especially given their persistent propensity to ignore new findings that do not confirm their earlier beliefs.

Conferences, Workshops, Seminars and Lectures

The following, incomplete, list includes presentations of the Principle Investigator of results of the MOE supported research at conferences, workshops, lectures and seminars. Essentially all were by invitation or by peer selection of proposed papers, and reflect the growing and wide international interest in the work. (MOE funds were not used for overseas travel.)

1988:

Assoc. Resource & Environ. Econ., New York -- Conference

Ont. Min. of the Environment, Toronto -- Conference

1989:

Western Regional Sci. Assoc., San Diego CA -- Conference

West Coast Law & Society Group, Vancouver -- Conference

Canadian Economics Association, Lavel Univ -- Conference

World Bank, Washington, U.S.A. -- Two day Workshop

Canadian Law & Economics Assoc., Toronto -- Conference

Banff Centre -- Five day Workshop

Ont. Min. of the Environment, Toronto -- Conference

1990:

Depart. of Fisheries and Oceans, Ottawa -- Seminar

Canadian Economics Assoc., Univ of Victoria -- Conference

Canadian Inst. of Resources Law, Calgary -- Conference

Banff Centre -- Five day Workshop

Hazardous Materials Facility Planning, Toronto -- Conference

1991:

Western Regional Science Assoc., Monterey CA -- Conference

Univer. of Alberta, Rural Economy -- Seminar

Univer. of Alberta, Business Faculty -- Seminar

Univer. of Lethbridge, Economics -- Seminar

Canadian Institute of Resources Law, Ottawa - - Conference

Canadian Economics Association, Queen's Univ -- Conference

Canadian Resource & Environ Econ., Guelph -- Conference

Banff Centre -- Five day Workshop

Ont. Min. of the Environment, Toronto -- Conference

Federation of ASEAN Econ. Assoc., Bangkok -- Conference

1992:

Ont. Resource Min., Sault St.Marie -- Four day workshop

Forestry and Environ, Jasper, Alberta -- Conference

Canadian Inst. of Resources Law, Calgary -- Conference

Canadian Economics Assoc., Univ of PEI -- Conference

Ont. Resource Min., Sault St.Marie -- Four day workshop

Law and Social Sci, Nat. Univ Singapore -- Conference

National University of Singapore, Economics -- Seminar

National University of Singapore -- Public Lecture

East Asian Econ. Assoc., Seoul, Korea -- Conference

University of Stirling, Economics, U.K. -- Seminar

University of Newcastle, Economics, U.K. -- Seminar

Univ. College London, Environ/Econ Centre -- Seminar

Banff Centre -- Five day Workshop

1993:

Australian Agr. Econ. Soc., Sydney, Australia -- Conference

Australian Resources Assessment Comm., Canberra -- Seminar

N.S.W., Environmental Prot. Authority, Sydney -- Conference

University of New England, Economics, Australia -- Seminar

Univ. of New England, Water Centre, Australia -- Seminar

B.C. Ministry of Environment, Victoria -- Conference

University of B.C., Law Faculty -- Seminar

Williams College, U.S.A., Economics -- Seminar

Williams College, U.S.A., Environment Center -- Seminar

Williams College, U.S.A., Intern'l Dev. Instit. -- Seminar

Colorado State Univer, U.S.A., Economics -- Seminar

University of Colorado, U.S.A., Economics -- Seminar

B.C. Government Agencies, Whistler, B.C. -- Conference

Soc. Advancement Behavioral Econ., NY, U.S.A. -- Conference

Papers and Publications

The papers produced from the research, and also identified as Attachments, include:

Attachment A
Knetsch, J.L., (1989), "The Endowment Effect and Evidence of Nonreversible Indifference Curves," The American Economic Review, 79:1277-84.

Attachment B -
Kahneman, Daniel, Jack L. Knetsch and Richard Thaler, (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem", Journal of Political Economy, 98:1325-48.

Attachment C -
Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler, (1991), "The Endowment Effect, Loss Aversion, and Status Quo Bias", Journal of Economics Perspectives, 5:193-206.

Attachment D -
Knetsch, J.L., (1992a), "Preferences and Nonreversibility of Indifference Curves," Journal of Economic Behavior and Organization, 17:131-139.

Attachment E -
Knetsch, J.L., (1993a), "Asymmetric Valuation of Gains and Losses and Preference Order Assumptions," Simon Fraser University Working Paper.

Attachment F -
Cohen, David, and Jack L. Knetsch, (forthcoming), "Judicial Choice and Disparities Between Measures of Economic Values", Osgoode Hall Law Journal.

Attachment G -
Borges, B., and J.L. Knetsch, (1993), "Valuation of Gains and Losses, Fairness, and Negotiation Outcomes", Working Paper, Simon Fraser University.

Attachment H -
Knetsch, J.L., (1990), "Environmental Policy Implications of Disparities Between Willingness to Pay and Compensation Demanded Measures of Values," Journal of Environmental Economics and Management, 18:227-237.

Attachment I -

Knetsch, Jack L., (1992b), "Economics, Losses, Fairness and Resource-Use Conflicts", in M. Ross and J.O. Saunders, eds., Growing Demands On A Shrinking Heritage: Managing Resource-Use Conflicts, Canadian Institute of Resources Law, Calgary.

Attachment J -
Knetsch, J.L., (1993c), "The Reference Point and Measures of Welfare Change", Working Paper, Simon Fraser University.

Attachment K -
Kahneman, Daniel, and Jack L. Knetsch, (1993), "A Test of the Double Referendum Contingent Valuation Question Format", Working Paper, Simon Fraser University.

Attachment L -
Kahneman, Daniel, and Jack L. Knetsch, (1992a), "Valuing Public Goods: The Purchase of Moral Satisfaction", Journal of Environmental Economics and Management, 22:57-70.

Attachment M -
Kahneman, Daniel, and Jack L. Knetsch, (1992b), "Contingent Valuation and the Value of Public Goods: Reply", Journal of Environmental Economics and Management, 22:90-94.

Attachment N -
Kahneman, Daniel, and Jack L. Knetsch, (1992c), "Issues of Method in the Study of Embedding: Reply to Harrison", Unpublished Working Paper, Simon Fraser University.

Behavioural Economics Course

While not a direct part of the research supported by MOE, the expanded research in the area of behavioural economics made possible by this support has indirectly led to the establishment of a new course in behavioural economics at Simon Fraser University. First offered as a "Selected Topics" course, it is now a permanently listed course offered annually, and draws undergraduate and graduate students from Economics, Business, and other disciplines. Many of the experiments, readings, and other materials used in the course are drawn from the MOE supported research.

References

Borges, B., and J.L. Knetsch, (1993), "Valuation of Gains and Losses, Fairness, and Negotiation Outcomes", Working Paper, Simon Fraser University.

Cambridge Economics, Inc., (1992), Proceedings of A Symposium Contingent Valuation: A Critical Assessment, Cambridge Economics, Inc.

Cohen, David, and Jack L. Knetsch, (forthcoming), "Judicial Choice and Disparities Between Measures of Economic Values", Osgoode Hall Law Journal.

Cummings, R.G., D.S. Brookshire, and W.D. Schulze, eds., (1986), Valuing Environmental Goods: An Assessment of the Contingent Valuation Method, Rowman and Allanheld.

Davis, Robert K., (1963), "The Value of Outdoor Recreation: An Economic Study of the Maine Woods" (unpublished PhD dissertation, Harvard University).

Desvousges, W.H., F.R. Johnson, R.W. Dunford, K.J. Boyle, S.P. Hudson, and K.N. Wilson, (1992), "Measuring Natural Resource Damages With Contingent Valuation: Tests of Validity and Reliability", in Cambridge Economics, Inc., Proceedings of A Symposium Contingent Valuation: A Critical Assessment, Cambridge Economics, Inc.

Edwards, Ward, and Detlof von Winterfeldt, (1987), Risk Analysis, 7:141-158.

Frey, Bruno, and Werner W. Pommerehne, (1987), "International Trade in Art: Attitudes and Behaviour", Rivista Internaxionale di Scienze Economiche a Commerciali, 34:465-486.

Gordon, Irene M, and Jack L. Knetsch, (1979), "Consumer's Surplus Measures and the Evaluation of Resources", Land Economics, 35:1-10.

Gregory, R., H. Kunreuther, D. Easterling, and K. Richards, (1990), "Incentive Policies to Site Hazardous Waste Facilities", Decision Research and University of Pennsylvania Working Paper.

Hammack, J., and Gardner Brown, (1974), Waterfowl and Wetlands: Toward Bioeconomic Analysis, Johns Hopkins Press.

Johnson, Eric J., John Hershey, Jacqueline Meszaros, and Howard Kunreuther, (forthcoming), Journal of Risk and Uncertainty.

Kachelmeier, S.J. and M. Shehata, (1992), "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China," The American Economic Review, 82:1120-1141

Kahneman, Daniel, and Jack L. Knetsch, (1992a), "Valuing Public Goods: The Purchase of Moral Satisfaction", Journal of Environmental Economics and Management, 22:57-70.

Kahneman, Daniel, and Jack L. Knetsch, (1992b), "Contingent Valuation and the Value of Public Goods: Reply", Journal of Environmental Economics and Management, 22:90-94.

Kahneman, Daniel, and Jack L. Knetsch, (1992c), "Issues of Method in the Study of Embedding: Reply to Harrison", Unpublished Working Paper, Simon Fraser University.

Kahneman, Daniel, and Jack L. Knetsch, (1993), "A Test of the Double Referendum Contingent Valuation Question Format", Working Paper, Simon Fraser University.

Kahneman, Daniel, Jack L. Knetsch and Richard Thaler, (1986a), "Fairness as a Constraint on Profit Seeking: Entitlements in the Market", The American Economic Review, 76:728-741.

Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler, (1986b), "Fairness and the Assumptions of Economics," Journal of Business, 59:S285-S300.

Kahneman, Daniel, Jack L. Knetsch and Richard Thaler, (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem", Journal of Political Economy, 98:1325-48.

Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler, (1991), "The Endowment Effect, Loss Aversion, and Status Quo Bias", Journal of Economics Perspectives, 5:193-206.

Kahneman, Daniel, and Ilana Ritov, (1993), "Determinants of Stated Willingness To Pay For Public Goods: A Study in the Headline Method", University of California Working Paper.

Kahneman, Daniel, and Amos Tversky, (1979), "Prospect Theory: An Analysis of Decisions Under Risk," Econometrica, 47:263-91.

Kealy, M.J., and R.W. Turner, (1993), "A Test of the Equality of Close Ended and Open Ended Contingent Valuations", American Journal of Agricultural Economics.

Knetsch, J.L., (1989), "The Endowment Effect and Evidence of Nonreversible Indifference Curves," The American Economic Review, 79:1277-84.

Knetsch, J.L., (1990), "Environmental Policy Implications of Disparities Between Willingness to Pay and Compensation Demanded Measures of Values," Journal of Environmental Economics and Management, 18:227-237.

Knetsch, J.L., (1992a), "Preferences and Nonreversibility of Indifference Curves," Journal of Economic Behavior and Organization, 17:131-139.

Knetsch, Jack L., (1992b), "Economics, Losses, Fairness and Resource-Use Conflicts", in M. Ross and J.O. Saunders, eds., Growing Demands On A Shrinking Heritage: Managing Resource-Use Conflicts, Canadian Institute of Resources Law, Calgary.

Knetsch, J.L., (1993a), "Asymmetric Valuation of Gains and Losses and Preference Order Assumptions," Simon Fraser University Working Paper.

Knetsch, J.L., (1993b), "The Reference Point and Measures of Welfare Change", Working Paper, Simon Fraser University.

Knetsch, J.L., and J.A. Sinden, (1984), "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," The Quarterly Journal of Economics, 99:507-521.

Mitchell, Robert Cameron, and Richard T. Carson, (1989), Using Surveys to Value Public Goods: The Contingent Valuation Method, Resources For The Future.

NOAA Panel, (1993), "Report of the NOAA Panel on Contingent Valuation", U.S. Federal Register, 58:4602-4614, (January 15).

Ontario Law Reform Commission, (1990), Report on Damages for Environmental Harm, Toronto.

Posner, Richard, (1986), Economic Analysis of Law, 3rd ed., Little, Brown and Company.

Randall, Alan, (1987), Resource Economics, 2nd ed., Wiley and Son.

Rhoads, S.E., (1985), The Economist's View of the World, Cambridge University Press.

Shefrin, Hersh, and Meir Statman, (1985), "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence", Journal of Finance, 40:777-790.

Simon, Herbert A., (1991), Models of My Life, Basic Books.

ATTACHMENT A

The Endowment Effect and Evidence of Nonreversible Indifference Curves

By JACK L. KNETSCH*

Indifference curves are normally taken to indicate corresponding tradeoffs of goods *A* for *B* or *B* for *A* over the same interval: "the rate of commodity substitution at a point on an indifference curve is the same for movements in either direction" (James Henderson and Richard Quandt, 1971, p. 12). However, recent empirical findings of asymmetric evaluations of gains and losses imply that the presumed reversibility may not accurately reflect preferences, and that people commonly make choices that differ depending on the direction of proposed trades.

The evidence from a wide variety of tests is consistent with the suggestions of Daniel Kahneman and Amos Tversky (1979), and Richard Thaler (1980), that losses from a reference position are systematically valued far more than commensurate gains. The minimum compensation people demand to give up a good has been found to be several times larger than the maximum amount they are willing to pay for a commensurate entitlement. For example, when questioned about the possible destruction of a duck habitat, hunters responded that they would be willing to pay an average of $247 to prevent its loss but would demand $1044 to accept it (Judd Hammack and Gardner Brown, 1974). Respondents in another study demanded payments to accept various levels of visibility degradation that were from 5 to over 16 times higher than their valuations based on payment measures (Robert Rowe, Ralph d'Arge, and David Brookshire, 1980). Similar large differences in valuations of a wide variety of goods and entitlements have been reported in numerous survey studies.[1]

These valuation differences based on responses to hypothetical survey questions prompted further, more severe, tests of the disparity using a variety of real exchange experimental designs. The results have been much the same. Even when exchanges of real goods and actual cash payments motivated the evaluations, the compensation demanded to give up an entitlement has been reported to far exceed the comparable payment measure of value (Knetsch and John Sinden, 1984; Thomas Heberlein and Richard Bishop, 1985).

While some studies have found that the size of an initial disparity may decrease over successive valuation trials for some entitlements, differences between the measures generally remain (Don Coursey, John Hovis, and William Schulze, 1987, and Knetsch and Sinden, 1987). For example, Brookshire and Coursey found that the differences between payment and compensation valuations of trees in a park decreased substantially from initial hypothetical responses, but the ratio of the two measures remained at "approximately five to one" after the last real exchange iteration (1987, p. 563). Other more recent tests of the evaluation disparities, or endowment effects, have shown that the observed disparities between valuations of entitlements are not the result of wealth effects or income constraints, strategic behavior, or transactions costs, and that the differences persist over repeated binding iterations of

---

*Simon Fraser University, Burnaby, Canada. This research was supported by Fisheries and Oceans Canada and The Ontario Ministry of the Environment; it has benefited from discussions with Curtis Eaton, Daniel Kahneman, Peter Kennedy, and Richard Thaler.

[1] See, for example, Ronald Cummings et al. (1986) for a summary of many earlier studies.

market trials (Knetsch, Thaler, and Kahneman, 1988).

The purpose here is to report the results of more direct tests of the reversibility of indifference curves. A variety of test designs were used in these exercises, but in each case one-half of a sample of participants were offered one good or money for another good, and the other half were offered a similar trade but in the opposite direction.[2]

I. Test 1: Exchanges of Two Goods

A straightforward preference exercise involving a choice between two goods was conducted with three comparable groups of University of Victoria students.

In one class, 76 participants were given a coffee mug and were then asked to complete a short questionnaire. The students had the mug in their possession while they answered the written questions. After the questionnaires were completed, the experimenter showed the participants a 400-gram Swiss chocolate bar and told them they could have one in exchange for their mug.[3] Participants were informed that they could either keep their mugs, or have a chocolate bar. They were instructed to hold up a colored paper with the word "trade" marked on it if they preferred the candy to their mug and wished to make an exchange. All desired trades were made immediately by one of four experimenters present, there was no uncertainty of receiving the other good and no effort was required beyond raising the paper to indicate a willingness to make an exchange.

The 87 participants in the second group were offered an opportunity to make the opposite trade of giving up a candy bar, which had been given to them initially, for a mug. All conditions were the same except for the direction of the exchange offer.

The 55 people in the third group were simply offered a choice between receiving a candy bar or mug.

The choices presented to people in each of the first two groups differed only in terms of which good they had to give up to obtain the other—incentives were compatible and there were no income, or wealth, effect possibilities, nor any wealth constraints. However, contrary to the expectation of an equal proportion favoring one good over the other in each group, based on the conventional assertion of economic theory and practice, the different initial entitlements and subsequent direction of potential trades heavily influenced the participants' valuations of the two goods.

The preferences indicated by the choices were:

	Proportion Favoring (In Percent)		
Group	Mug Over Candy	Candy Over Mug	*N*
1. Give up mug to obtain candy	89	11	76
2. Give up candy to obtain mug	10	90	87
3. No initial entitlement	56	44	55

When given a simple choice without a prior entitlement or reference position, 56 percent of the participants (Group 3) selected a mug in preference to a candy bar. But only 10 percent valued a mug more than a candy bar when they had to give up the candy to obtain a mug (Group 2), while 89 percent of those who initially received a mug (Group 1) declined to give it up for a candy bar. The preferences for a mug over a candy bar varied from 10 to 89 percent depending entirely on the reference position of the endowment at the time of the valuations.

In this simple experiment, participants' preferences were not independent of the direction of the exchanges, as is commonly assumed. They expressed a dramatic asymmetry in valuations by weighing the loss of giving up their initial or reference entitlement far more heavily than the foregone gains of not obtaining the alternative entitlement.

---

[2] Instructions for all of the exercises are available from the author.

[3] Participants were familiar with both the mugs and the candy bars. The mugs were sold in the University bookstore for $4.95 and the candy bars were available in local shops for around $6.

II. Test 2: Minimum Acceptable Exchanges

The symmetry of exchange preferences was further tested in a second experiment involving real goods and actual cash payments. As in the first test, the design of this experiment also precluded the influence of a wealth constraint and the possibility of an income, or wealth, effect. The participants were 80 students in five tutorial groups of an introductory pre-principles economics class at Simon Fraser University. The exercise was repeated in the five tutorial meetings in a single afternoon.

Forty-one individuals in three of the groups were first given two 100-gram Swiss chocolate candy bars.[4] They were then asked to state the smallest amount of money they would accept to give up both of their candy bars.

The 39 members of the other two tutorial groups were each given two single-dollar bills. They were then asked the minimum number of chocolate bars they would require to give up their two dollars. The initial offering of money or candy bars was alternated over the five participating groups.

The method of obtaining the declarations of minimum acceptable trades was designed so that stating the smallest compensation each participant would agree to was in the person's best interest. To ensure that the individuals understood that this was the case, a preliminary instruction and screening test was conducted so that they could demonstrate their understanding before becoming participants in the real exchange exercise. This preliminary test involved a hypothetical, induced, or prescribed, value and the identical bidding mechanism that was used to reveal preferences in the subsequent real good exchange.[5]

All participants were given the same form for the preliminary demonstration and screening test:

> THIS FIRST PART OF THIS EXERCISE IS HYPOTHETICAL ONLY; NO ACTUAL PAYMENTS WILL BE MADE. HOWEVER, DEPENDING ON THE RESPONSE OF PARTICIPANTS IN THIS GROUP, A SECOND PART MAY BE CONDUCTED IN WHICH ACTUAL PAYMENTS WILL BE MADE.
>
> IF THE SECOND PART IS CONDUCTED, YOU CAN ONLY WIN, YOU CANNOT LOSE.
>
> Assume that the person conducting this exercise will pay you $2.50 for the ticket attached to this sheet.
>
> Now assume that before you redeem the ticket you are given an opportunity to sell it. The price you will be offered for it will be determined by a random draw of one of six cards with the dollar amounts: $0, $1, $2, $3, $4, or $5 on them.
>
> What is the smallest of these offer prices you would accept to give up your ticket?
>
> $______ (use only whole dollars—i.e., no cents). If the offer price on the randomly drawn card is the same or higher than the price you stated above, you will receive the amount of money on the card in exchange for your ticket. If the offer price on the card is lower

---

[4] These candy bars sell for around $2.00 each in most shops.

[5] Real exchange experiments using such prescribed, or induced, values have been widely used to test the efficiency of various types of market institutions (for example, Vernon Smith, 1976).

A comparison of the outcomes of induced value markets with subsequent goods markets carried out with identical trading rules also provides a test for possible transactions costs and strategic behavior as possible sources of valuation disparities (Knetsch, Thaler, and Kahneman, 1988). If no inhibitions to trades are observed in the induced-value markets, none should then be expected from such causes in markets for goods that are valued by the individual participants themselves.

> than your indicated minimum selling price, you will keep the ticket (and turn it in for the $2.50).

After the participants read the instructions for the preliminary test, the experimenter gave a further explanation and answered questions. Participants then completed the forms. The experimenter checked the forms to ensure that at least 75 percent of the individuals knew how to respond in a way that served their best interest (by stating $3 as the minimum acceptable price). If more than 25 percent failed to answer in this way the experiment was terminated for that whole group.[6]

Of the 80 individuals in the five groups that went on to the second part involving the real exchanges, only 14 (18 percent) failed to answer the first part correctly.[7] Before proceeding with the second part, the experimenter explained the bidding mechanism again, and pointed out the reasons why other answers could work to the disadvantage of the respondent. These further explanations, and the screening of participants by their demonstrated ability to reveal their true minimum values in this market, ensured that individuals participating in the real exchange experiment fully understood the value revealing incentives of the bidding procedure.

Participants in the second, real exchange, part of the experiment were given their endowment of either candy bars or dollar bills and were told that these were theirs to keep. Those with an entitlement of candy bars then received the following second form:

> THIS SECOND PART WILL INVOLVE REAL PAYMENTS.
> THERE ARE NO RIGHT OR WRONG ANSWERS, YOUR RESPONSE SHOULD REFLECT ONLY YOUR OWN PREFERENCES.
>
> You have been given 2 candy bars. They are yours to keep. However, if you like—and the choice is entirely yours—you may be able to trade the 2 candy bars for money. We want to know about your personal preferences in terms of the smallest number of dollars you would accept to give up the two candy bars.
>
> The amount you will be offered to give up your 2 candy bars will be determined in the same way as it was in Part 1 of this exercise. That is, an offer price, in dollars, will be determined by a random draw of one of the six cards with amounts ranging from $0 to $5.
>
> If you are willing to give up the 2 candy bars for $5 or less, write the smallest of these offer prices that you would accept to give up both of the candy bars in the blank below.
>
> If it would take more than $5 for you to agree to give up your two candy bars, then write a 6 in the blank, and you will then keep your candy bars.
>
> $______ (use only whole dollars—i.e., no cents). If the offer amount on the randomly drawn card is the same or higher than the number you have written above, you will receive the amount of money on the card in exchange for your 2 candy bars. If the offer on the card is lower than the number you have written above you will be allowed to keep the candy bars.

Individuals receiving an initial endowment of money were given a comparable form.

The bid procedure was then explained yet again and questions answered. The random draw of the card was made after the forms were completed, and exchanges were made according to the indicated minimum acceptable trades.

The results again demonstrated a strong aversion to giving up an initial entitlement.

The participants' dollar value of the candy bars is given either by the minimum amount demanded to give them up, or by the number of candy bars required to give up dollars. The proportions of the two samples valuing a candy bar equal to or greater than $1 and $2 are given in the following tabulation.

---

[6] This occurred in four other groups.

[7] All but 3 of the 14 individuals answering incorrectly specified a selling price *lower* than the most profitable one, an error that would bias the results toward lower evaluations of the initial entitlements and symmetry and away from the much higher values of losses subsequently revealed.

Group	Proportion of Individuals Valuing Candy Bar Equal to or More than: (In Percent) $1	$2
1. Give up Money to Get Candy Bars ($N=39$)	33	8
2. Give Up Candy Bars to Get Money ($N=41$)	95	37
Chi-Squared Value	35.05	9.57

The participants valued the candy or money significantly more when considering the prospect of giving it up than they did when they had the opportunity to acquire either in exchange for their original entitlement.[8] The average evaluations of the candy bars, based on the responses of all of the participants in each original endowment group, were approximately $1.83 with the entitlement and $0.90 without it. That is, the people starting with the candy bars valued them at a little over twice as much as the people did who were given money.[9]

As in the earlier test, the relative preferences of participants in this exercise depended on reference entitlements and the subsequent direction of trade offers.

III. Test 3: Survey Responses

While the results from the first two tests were from real exchange experiments, similar indications of large valuation disparities that varied with the direction of proposed exchanges were also obtained from general public reactions recorded in a series of telephone surveys.

These random household telephone interviews were conducted in metropolitan Toronto, with an equal number of adult males and females called during early evening hours. In each case, one subsample of respondents was asked if they would trade one good for another; a second subsample was asked if they would make the opposite exchange.

Two groups of respondents were asked about their preferences between a change of $700 per year in their current income and a 0.5 percent change in the chance of having an accident during the year that would result in some time in the hospital. Among respondents in the group asked if they would accept $700 to have their chance of an accident increased by 0.5 percent (from 0.5 percent to 1.0 percent), 61 percent ($N=149$) refused, indicating that this proportion valued the negative change in risk more than $700.

In the group asked to trade a decrease of the same accident risk (over the same range of from 1.0 percent to 0.5 percent) for a reduction of $700 in their income, only 27 percent ($N=146$) indicated a higher valuation of the positive increment of risk change by being willing to make the trade. Again, relative preferences varied greatly depend-

[8] Based on 1 degree of freedom, both chi-squared values are significant at the 1 percent level.

[9] The differences indicated in the tabulation and the statistical significance of the disparity are, if anything, understated. First, for people stating minimum money demands the recorded numbers are those valuing the money *more than* the candy. Individuals in the other groups stating required candy bar payments include those who value the candy *equal to* as well as more than the various sums.

Second, the proportions of participants valuing the bars at the stated sums imply that the individuals valued all candy bars and dollars over the relevant range equally. If they in fact valued each successive candy bar or dollar by a decreasing amount—with diminishing marginal utility—then the surplus from the intramarginal units would prompt an even greater demand for the other good, given the all or nothing nature of the choice.

ing on the direction of the proposed exchange—39 percent preferred $700 to the change in risk one way, and 73 percent preferred $700 to the change the other way.[10]

In the same way, two groups of respondents were asked about their preferences between a new job that offered more or less pay and more or less vacation time than their current employment. When asked to trade an increase of $500 annual salary for one week less vacation, 66 percent ($N=143$) indicated a greater valuation of the holiday time. For respondents offered the opposite exchange of more vacation for less pay, only 29 percent ($N=147$) indicated a preference for an equal amount of time off.[11]

IV. The Asymmetry of Preferences

The evidence from these three tests is widely at variance with the conventional strong empirical assertion of completely reversible indifference curves. The tests used three different designs which varied in several dimensions: they included both real exchange experiments and surveys using hypothetical questions; they used student and nonstudent participants; they involved choices between goods and money and between two goods; they included small and large sums; and they included valuations that could be affected by wealth effects and income restraints and others which excluded such possibilities. The results, however, were all the same. In every case, people exhibited valuations and preferences that varied systematically and substantially with the initial reference entitlement and the direction of exchange offers.

Further, the windfall nature of the endowment in the real exchange tests could well lead to these results understating the extent of the disparities that might be expected under nonexperimental conditions. People were given a good or money in these experiments, and their valuations were assessed in terms of what they would demand to give it up. However, individuals have generally been found to give up such windfall gains more easily than assets from their own original endowment (Thaler and Eric Johnson, 1989). The manipulations used in these experiments may therefore be weaker than might be encountered in most actual assessments—and indifference curves may be even less reversible.

The results of the exercises reported here are consistent with other empirical tests of people's evaluations. The greater value ascribed to entitlements that might be lost over ones that might be gained gives rise to the disparity between bid and offer values (Kahneman and Tversky, 1979 and 1984). This appears to be similar to the observed nonreversibility of indifference curves in the data reported here.

Substantial disparities between valuations have now been reported for a number of goods, including an array of environmental assets, lottery tickets, personal safety, and many consumer goods. Except for cases of transparent choices, such as the obvious worth of alternatives presented in induced-value experimental markets where specific values are assigned by experimenters and individuals act on the basis of these valuations, there is yet little indication of the limits to the range of goods and assets where such differences might not be found. This raises the possibility that the significant endowment effects and nonreversible indifference curves found in these exercises may exist in a larger class of cases. Only additional empirical study can determine the extent that this is the case.

If the existence of nonreversible indifference curves is pervasive, then the assumption of reversibility will represent a normative ideal but will not serve as a very useful description of actual economic behavior in most circumstances. Nor will the traditional assumption of symmetry of valuations and complete reversibility of indifference curves seem as useful a basis for predicting economic behavior or as defendable a basis for the common practice of substituting payment measures for compensation measures of welfare losses (Robert Willig, 1976). The common practice of assessing losses in economic well-being by people's willingness to pay to avoid the change may result in significant understatements if indifference curves are not reversible.

---

[10] The chi-squared value of 35.32 with 1 d.f. is significant at the 0.1 percent level.

[11] With 1 d.f. the chi-squared value of 43.77 is significant at a 0.1 percent level.

The widespread irreversibility of indifference curves would not imply that people will not make any trades, or that consumers will not change future consumption patterns in response to changing relative prices. However, the presence of irreversibilities would imply that fewer trades will be made than predicted by standard assumptions, and they offer little assurance that the shifts in future consumption will be as complete or as prompt as would be expected if indifference curves were reversible.

In a similar way, the presence of irreversibilities also suggests that common presumptions of the potential gains from trade may often be overstated. In terms of an Edgeworth box, the area representing mutually advantageous exchanges may be very much smaller, or may not even exist, if curves showing points of indifference when giving up initial allocations differ from those indicating trades in the other direction. An apparent reluctance to trade may be due to the availability of few real gains to the parties. And the commonly observed aversion to giving up concessions in negotiations may well be due to a similar asymmetry in valuations of gains and losses (Kahneman, Knetsch, and Thaler, 1986).

REFERENCES

Brookshire, David S. and Coursey, Don L., "Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures," *American Economic Review*, September 1987, *77*, 554–66.

Coursey, Don L., Hovis, John J. and Schulze, William D., "The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value," *Quarterly Journal of Economics*, August 1987, *102*, 679–90.

Cummings, R. G., Brookshire, D. S. and Schulze, W. D., eds., *Valuing Environmental Goods*, Totowa, NJ: Rowman and Allanheld, 1986.

Hammack, J. and Brown, Gardner, *Waterfowl and Wetlands: Toward Bioeconomic Analysis*, Baltimore, MD: Johns Hopkins Press, 1974.

Heberlein, Thomas A. and Bishop, Richard C., "Assessing the Validity of Contingent Valuation: Three Field Experiments," Paper presented at the *International Conference on Man's Role in Changing the Global Environment*, Italy, 1985.

Henderson, James M. and Quandt, Richard E., *Microeconomic Theory*, 2nd ed., New York: McGraw-Hill, 1971.

Kahneman, Daniel and Tversky, Amos, "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrica*, March 1979, *47*, 263–91.

______ and ______, "Choices, Values, and Frames," *American Psychologist*, April 1984, *39*, 341–50.

______, Knetsch, Jack L. and Thaler, Richard, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, September 1986, *76*, 728–41.

Knetsch, Jack L., Thaler, Richard and Kahneman, Daniel, "Experimental Tests of the Endowment Effect and the Coase Theorem," Simon Fraser University Working Paper, 1988.

______ and Sinden, J. A., "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," *Quarterly Journal of Economics*, August 1984, *99*, 507–21.

______ and ______, "The Persistence of Evaluation Disparities," *Quarterly Journal of Economics*, August 1987, *102*, 691–95.

Rowe, Robert D., d'Arge, Ralph C. and Brookshire, David S., "An Experiment on the Economic Value of Visibility," *Journal of Environmental Economics and Management*, March 1980, *7*, 1–19.

Smith, Vernon L., "Experimental Economics: Induced Value Theory," *American Economic Review*, May 1976, *66*, 274–79.

Thaler, Richard, "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization*, March 1980, *1*, 39–60.

______ **and Johnson, Eric J.,** "The Effect of Prior Outcomes on Risky Choice," Cornell University Working Paper, 1989.

Willig, Robert D., "Consumer's Surplus Without Apology," *American Economic Review*, September 1976, *66*, 589–97.

ATTACHMENT B

Experimental Tests of the Endowment Effect and the Coase Theorem

---

Daniel Kahneman

University of California, Berkeley

Jack L. Knetsch[*]

Simon Fraser University

Richard H. Thaler

Cornell University

Contrary to theoretical expectations, measures of willingness to accept greatly exceed measures of willingness to pay. This paper reports several experiments that demonstrate that this "endowment effect" persists even in market settings with opportunities to learn. Consumption objects (e.g., coffee mugs) are randomly given to half the subjects in an experiment. Markets for the mugs are then conducted. The Coase theorem predicts that about half the mugs will trade, but observed volume is always significantly less. When markets for "induced-value" tokens are conducted, the predicted volume is observed, suggesting that transactions costs cannot explain the undertrading for consumption goods.

I. Introduction

The standard assumptions of economic theory imply that when income effects are small, differences between an individual's maximum

---

[*] Financial support was provided by Fisheries and Oceans Canada, the Ontario Ministry of the Environment, and the behavioral economics program of the Alfred P. Sloan Foundation. We wish to thank Vernon Smith for encouraging us to conduct these experiments and for providing extensive comments on earlier drafts. Of course, the usual disclaimer applies.

[*Journal of Political Economy*, 1990, vol. 98, no. 6]
© 1990 by The University of Chicago. All rights reserved. 0022-3808/90/9806-0004$01.50

willingness to pay (WTP) for a good and minimum compensation demanded for the same entitlement (willingness to accept [WTA]) should be negligible (Willig 1976). Thus indifference curves are drawn without reference to current endowments; any difference between equivalent and compensating variation assessments of welfare changes is in practice ignored;[1] and there is wide acceptance of the Coase theorem assertion that, subject to income effects, the allocation of resources will be independent of the assignment of property rights when costless trades are possible.

The assumption that entitlements do not affect value contrasts sharply with empirical observations of significantly higher selling than buying prices. For example, Thaler (1980) found that the minimal compensation demanded for accepting a .001 risk of sudden death was higher by one or two orders of magnitude than the amount people were willing to pay to eliminate an identical existing risk. Other examples of similar reported findings are summarized in table 1. The disparities observed in these examples are clearly too large to be explained plausibly by income effects.

Several factors probably contribute to the discrepancies between the evaluations of buyers and sellers that are documented in table 1. The perceived illegitimacy of the transaction may, for example, contribute to the extraordinarily high demand for personal compensation for agreeing to the loss of a public good (e.g., Rowe, d'Arge, and Brookshire 1980). Standard bargaining habits may also contribute to a discrepancy between the stated reservation prices of buyers and sellers. Sellers are often rewarded for overstating their true value, and buyers for understating theirs (Knez, Smith, and Williams 1985). By force of habit they may misrepresent their true valuations even when such misrepresentation confers no advantage, as in answering hypothetical questions or one-shot or single transactions. In such situations the buying-selling discrepancy is simply a strategic mistake, which experienced traders will learn to avoid (Coursey, Hovis, and Schulze 1987; Brookshire and Coursey 1987).

The hypothesis of interest here is that many discrepancies between WTA and WTP, far from being a mistake, reflect a genuine effect of reference positions on preferences. Thaler (1980) labeled the increased value of a good to an individual when the good becomes part of the individual's endowment the "endowment effect." This effect is a manifestation of "loss aversion," the generalization that losses are weighted substantially more than objectively commensurate gains in

---

[1] For example, the conventional prescription for assessing environmental and other losses is that, "practically speaking, it does not appear to make much difference which definition is accepted" (Freeman 1979, p. 3).

TABLE I

SUMMARY OF PAST TESTS OF EVALUATION DISPARITY

	MEANS			MEDIANS		
STUDY AND ENTITLEMENT	WTP	WTA	Ratio	WTP	WTA	Ratio
Hypothetical surveys:						
Hammack and Brown (1974): marshes	$247	$1,044	4.2			
Sinclair (1978): fishing				35	100	2.9
Banford et al. (1979):						
Fishing pier	43	120	2.8	47	129	2.7
Postal service	22	93	4.2	22	106	4.8
Bishop and Heberlein (1979): goose hunting permits	21	101	4.8			
Rowe et al. (1980): visibility	1.33	3.49	2.6			
Brookshire et al. (1980): elk hunting*	54	143	2.6			
Heberlein and Bishop (1985): deer hunting	31	513	16.5			
Real exchange experiments:						
Knetsch and Sinden (1984): lottery tickets	1.28	5.18	4.0			
Heberlein and Bishop (1985): deer hunting	25	172	6.9			
Coursey et al. (1987): taste of sucrose octa-acetate†	3.45	4.71	1.4	1.33	3.49	2.6
Brookshire and Coursey (1987): park trees†‡	10.12	56.60	5.6	6.30	12.96	2.1

* Middle-level change of several used in study.

† Final values after multiple iterations.

‡ Average of two levels of tree plantings.

the evaluation of prospects and trades (Kahneman and Tversky 1979; Tversky and Kahneman, in press). An implication of this asymmetry is that if a good is evaluated as a loss when it is given up and as a gain when it is acquired, loss aversion will, on average, induce a higher dollar value for owners than for potential buyers, reducing the set of mutually acceptable trades.

There are some cases in which no endowment effect would be expected, such as when goods are purchased for resale rather than for utilization. A particularly clear case of a good held exclusively for resale is the notional token typically traded in experimental markets commonly used to test the efficiency of market institutions (Plott 1982; Smith 1982). Such experiments employ the induced-value technique in which the objects of trade are tokens to which private redemption values that vary among individual participants have been assigned by the experimenter (Smith 1976). Subjects can obtain the prescribed value assigned for the tokens when redeeming them at the end of the trading period; the tokens are otherwise worthless.

No endowment effect would be expected for such tokens, which are valued only because they can be redeemed for cash. Thus both buyers and sellers should value tokens at the induced value they have been assigned. Markets for induced-value tokens can therefore be used as a control condition to determine whether differences between the values of buyers and sellers in other markets could be attributable to transaction costs, misunderstandings, or habitual strategies of bargaining. Any discrepancy between the buying and selling values can be isolated in an experiment by comparing the outcomes of markets for real goods with those of otherwise identical markets for induced-value tokens. If no differences in values are observed for the induced-value tokens, then economic theory predicts that no differences between buying and selling values will be observed for consumption goods evaluated and traded under the same conditions.

The results from a series of experiments involving real exchanges of tokens and of various consumption goods are reported in this paper. In each case, a random allocation design was used to test for the presence of an endowment effect. Half of the subjects were endowed with a good and became potential sellers in each market; the other half of the subjects were potential buyers. Conventional economic analysis yields the simple prediction that one-half of the goods should be traded in voluntary exchanges. If value is unaffected by ownership, then the distribution of values in the two groups should be the same except for sampling variation. The supply and demand curves should therefore be mirror images of each other, intersecting at their common median. The null hypothesis is, therefore, that half of the goods provided should change hands. Label this predicted volume V^*. If there is an endowment effect, the value of the good will be higher for sellers than for buyers, and observed volume V will be less than V^*. The ratio V/V^* provides a unit-free measure of the undertrading that is produced by the effect of ownership on value. To test the hypothesis that market experience eliminates undertrading, the markets were repeated several times.

A test for the possibility that observed undertrading was due to transaction costs was provided by a comparison of the results from a series of induced-value markets with those from the subsequent goods markets carried out with identical trading rules. Notice that this comparison can also be used to eliminate numerous other possible explanations of the observed undertrading. For example, if the instructions to the subjects are confusing or misleading, the effects should show up in both the induced-value markets and the experimental markets for real goods. Section II describes studies of trading volume in induced-value markets and in consumption goods markets. Section III provides a further test for strategic behavior and demonstrates that the disparity findings are not likely caused by this. Section IV investigates the extent to which the undertrading of goods is produced by reluctance to buy and reluctance to sell. Section V examines undertrading in bilateral negotiations and provides a test of the Coase theorem. Section VI describes an experiment that rules out income effects and a trophy effect as explanations of the observed valuation disparity. Implications of the observed effects are discussed in Section VII.

II. Repeated Market Experiments

In experiment 1, 44 students in an advanced undergraduate law and economics class at Cornell University received a packet of general instructions plus 11 forms, one for each of the markets that were conducted in the experiment. (The instructions for all experiments are available from the authors.) The first three markets were conducted for induced-value tokens. Sellers received the following instructions (with differences for buyers in brackets):

> In this market the objects being traded are tokens. You are an owner, so you now own a token [You are a buyer, so you have an opportunity to buy a token] which has a value to you of \$x. It has this value to you because the experimenter will give you this much money for it. The value of the token is different for different individuals. A price for the tokens will be determined later. For each of the prices listed below, please indicate whether you prefer to: (1) Sell your token at

> this price and receive the market price. [Buy a token at this price and cash it in for the sum of money indicated above.] (2) Keep your token and cash it in for the sum of money indicated above. [Not buy a token at this price.] For each price indicate your decision by marking an X in the appropriate column.

Part of the response form for sellers follows:

At a price of \$8.75 I will sell _____ I will not sell _____

At a price of \$8.25 I will sell _____ I will not sell _____

The same rectangular distribution of values—ranging from \$0.25 to \$8.75 in steps of \$0.50—was prepared for both buyers and sellers. Because not all the forms were actually distributed, however, the induced supply and demand curves were not always precisely symmetrical. Subjects alternated between the buyer and seller role in the three successive markets and were assigned a different individual redemption value in each trial.

Experimenters collected the forms from all participants after each market period and immediately calculated and announced the market-clearing price,[2] the number of trades, and the presence or absence of excess demand or supply at the market-clearing price.[3] Three buyers and three sellers were selected at random after each of the induced markets and were paid off according to the preferences stated on their forms and the market-clearing price for that period.

Immediately after the three induced-value markets, subjects on alternating seats were given Cornell coffee mugs, which sell for \$6.00 each at the bookstore. The experimenter asked all participants to examine a mug, either their own or their neighbor's. The experimenter then informed the subjects that four markets for mugs would be conducted using the same procedures as the prior induced markets with two exceptions: (1) One of the four market trials would subsequently be selected at random, and only the trades made on this trial would be executed. (2) In the binding market trial, *all* trades would be implemented, unlike the subset implemented in the induced-value markets.[4] The initial assignment of buyer and seller roles was maintained for all four trading periods. The clearing price and the number of trades were announced after each period. The market that "counted" was indicated after the fourth period, and transactions were executed immediately. All sellers who had indicated that they would give up their mugs for a sum at the market-clearing price exchanged their mugs for cash, and successful buyers paid this same price and received their mugs. This design was used to permit learning to take place over successive trials and yet make each trial potentially binding. The same procedure was then followed for four more successive markets using boxed ballpoint pens with a visible bookstore price tag of $3.98, which were distributed to the subjects who had been buyers in the mug markets.

---

[2] The instructions stated that "*it is in your best interest to answer these questions truthfully.* For any question, treat the price as fixed. (In economics jargon, you should act as 'price takers'.)" All the subjects were junior and senior economics majors, so they were familiar with the terms used. If subjects asked how the market prices were determined, they were told, truthfully, that the market price was the point at which the elicited supply and demand curves intersected. The uniformity of the results across many different experiments suggests that this information had no discernible effect on behavior. Furthermore, the responses of the subjects in the induced-value portion of the experiments indicate that nearly all understood and accepted their role as price takers. See also experiment 5, in which a random price procedure was used.

[3] When this occurred, a random draw determined which buyers and sellers were accommodated.

For each goods market, subjects completed a form similar to that used for the induced-value tokens, with the following instructions:

> You now own the object in your possession. [You do not own the object that you see in the possession of some of your neighbors.] You have the option of selling it [buying one] if a price, which will be determined later, is acceptable to you. For each of the possible prices below indicate whether you wish to: (1) Sell your object and receive this price [Pay this price and receive an object to take home with you] or (2) Keep your object and take it home with you. [Not buy an object at this price.] For each price indicate your decision by marking an X in the appropriate column.

The buyers and sellers in the consumption goods markets faced the same incentives that they had experienced in the induced-value markets. Buyers maximized their potential gain by agreeing to buy at all prices below the value they ascribed to the good, and sellers maximized their welfare by agreeing to sell at all prices above the good's worth to them. As in the induced-value markets, it was in the best interest of the participants to act as price takers.

As shown in table 2, the markets for induced-value tokens and consumption goods yielded sharply different results. In the induced-value markets, as expected, the median buying and selling prices were identical. The ratio of actual to predicted volume (V/V^*) was 1.0,

---

[4] The experimental design was intended to give the markets for consumption goods every possible chance to be efficient. While in the induced-value markets not everyone was paid, in the consumption goods markets everyone was paid. Also, the consumption goods markets were conducted after the induced-value markets and were repeated four times each, to allow the subjects the maximum opportunity for learning.

TABLE 2

RESULTS OF EXPERIMENT 1

INDUCED-VALUE MARKETS

Trial	Actual Trades	Expected Trades	Price	Expected Price
1	12	11	3.75	3.75
2	11	11	4.75	4.75
3	10	11	4.25	4.25

CONSUMPTION GOODS MARKETS

Trial	Trades	Price	Median Buyer Reservation Price	Median Seller Reservation Price
		Mugs (Expected Trades = 11)		
4	4	4.25	2.75	5.25
5	1	4.75	2.25	5.25
6	2	4.50	2.25	5.25
7	2	4.25	2.25	5.25
		Pens (Expected Trades = 11)		
8	4	1.25	.75	2.50
9	5	1.25	.75	1.75
10	4	1.25	.75	2.25
11	5	1.25	.75	1.75

aggregating over the three periods. In contrast, the median selling prices in the mug and pen markets were more than twice the median buying prices, and the V/V^* ratio was only .20 for mugs and .41 for pens. Observed volume did not increase over successive periods in either the mug or the pen markets, providing no indication that subjects learned to adopt equal buying and selling prices.

The results of the first and last markets for coffee mugs are also displayed in figure 1. There are five features to notice in this figure: (1) Both buyers and sellers display a wide range of values, indicating that in the absence of an endowment effect there would be enough rents to produce gains from trade. Indeed, the range of values is similar to that used in the induced-value markets, which had near-perfect market efficiency. (2) The distribution of selling prices has a single mode, unlike some recent results in which an evaluation discrepancy could be explained by a bimodal distribution of compensation demanded (Boyce et al. 1990). (3) The payment of a small commission for trading, such as $0.25 per trade, would not significantly alter the results. (4) The mugs were desirable. Every subject assigned a positive value to the mug, and the lowest value assigned by a seller

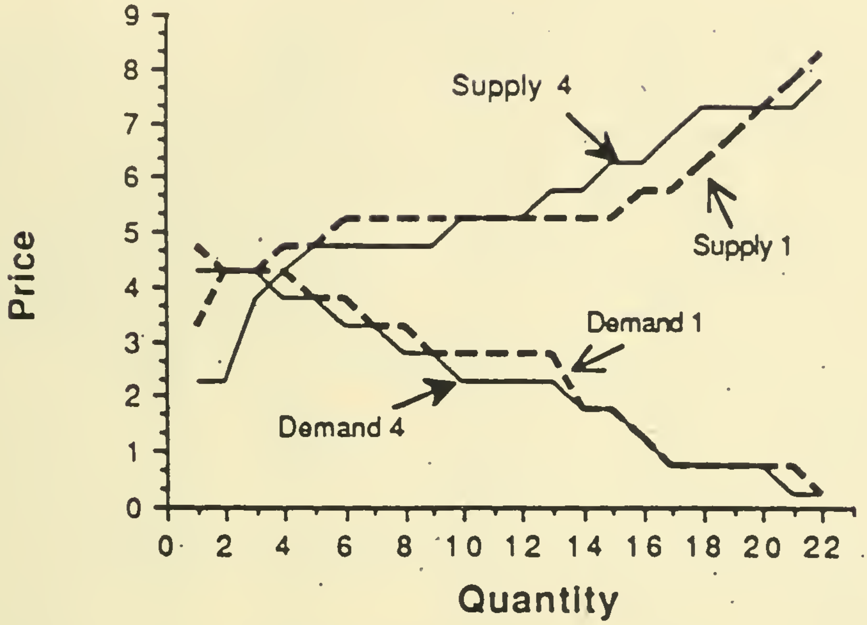


FIG. 1.—Supply and demand curves, markets 1 and 4

was $2.25. (5) Neither demand nor supply changed much between the first and last markets.

Experiment 2 was conducted in an undergraduate microeconomics class at Cornell ($N = 38$). The procedure was identical to that of experiment 1, except that the second consumption good was a pair of folding binoculars in a cardboard frame, available at the bookstore for $4.00. The results are reported in table 3.

In experiments 3 and 4, conducted in Simon Fraser University undergraduate economics classes, the subjects were asked to provide minimum selling prices or maximum buying prices rather than to answer the series of yes or no questions used in experiments 1 and 2. The induced-value markets were conducted with no monetary payoffs and were followed by four markets for pens in experiment 3 and five markets for mugs in experiment 4. In experiment 3, subjects were told that the first three markets for pens would be used for practice, so only the fourth and final market would be binding. In experiment 4, one of the five markets was selected at random to count, as in experiments 1 and 2. Other procedures were unchanged. The results are shown in table 4.

Experiments 2–4 all yielded results similar to those obtained in experiment 1. Summing over the induced-value markets in all four experiments produced a V/V^* index of .91. This excellent performance was achieved even though the participants did not have the

TABLE 3

Results of Experiment 2

Induced-Value Markets

Trial	Actual Trades	Expected Trades	Price	Expected Price
1	10	10	3.75	4.25
2	9	10	4.75	4.25
3	7	8	4.25	4.75

Consumption Goods Markets

Trial	Trades	Price	Median Buyer Reservation Price	Median Seller Reservation Price
	Mugs (Expected Trades = 9.5)			
4	3	3.75	1.75	4.75
5	3	3.25	2.25	4.75
6	2	3.25	2.25	4.75
7	2	3.25	2.25	4.25
	Binoculars (Expected Trades = 9.5)			
8	4	1.25	.75	1.25
9	4	.75	.75	1.25
10	3	.75	.75	1.75
11	3	.75	.75	1.75

benefit of experience with the trading rules, there were limited monetary incentives in experiments 1 and 2, and there were no monetary incentives in experiments 3 and 4. In the markets for consumption goods, in which all participants faced monetary incentives and experience with the market rules gained from the induced-value markets, V/V^* averaged .31, and median selling prices were more than double the corresponding buying prices. Trading procedures were precisely identical in markets for goods and for induced-value tokens. The high volume of trade in money tokens therefore eliminates transaction costs (or any other feature that was present in both types of markets) as an explanation of the observed undertrading of consumption goods.

It should be noted that subjects in the position of buyers were not given money to use for purchases, but rather had to make transactions using their own money. (Subjects were told to bring money to class and that credit and change would be available if necessary. Some subjects borrowed from friends to make payments.) The aim was to study transactions in a realistic setting. While the present design makes potential sellers slightly wealthier, at least in the first market, the magnitude of the possible income effect is trivial. In one of the

TABLE 4

Results of Experiments 3 and 4

Trial	N	Object	Actual Trades	Expected Trades	Ratio of Seller Median Value to Buyer Median Value
			Experiment 3		
1	26	Induced	5	6.5	
2	26	Pen	2	6.5	6.0
3	26	Pen	2	6.5	6.0
4	26	Pen	2	6.5	5.0
5	26	Pen	1	6.5	5.0
			Experiment 4		
1	74	Induced	15	18.5	
2	74	Induced	16	18.5	
3	74	Mug	6	18.5	3.8
4	74	Mug	4	18.5	2.8
5	72	Mug	4	18	2.2
6	73	Mug	8	18	1.8
7	74	Mug	8	18.5	1.8

markets the equilibrium price was only $0.75, and the prices in other markets were never above a few dollars. Also, as shown in experiments 7 and 8 below, equal undertrading was found in designs that eliminated the possibility of an income effect or cash constraint.

As shown in tables 2–4, subjects showed almost no undertrading even in their first trial in an induced-value market. Evidently neither bargaining habits nor any transaction costs impede trading in money tokens. On the other hand, there is no indication that participants in the markets for goods learned to make valuations independent of their entitlements. The discrepant evaluations of buyers and sellers remained stable over four, and in one case five, successive markets for the same good and did not change systematically over repeated markets for successive goods.

A difference in procedure probably explains the apparent conflict between these results and the conclusion reached in some other studies, that the WTA-WTP discrepancy is greatly reduced by market experience. The studies that reported a disciplinary effect of market experience assessed this effect by comparing the responses of buyers and sellers in preliminary hypothetical questions or nonbinding market trials to their behavior in a subsequent binding trial with real monetary payoffs (Knez et al. 1985; Brookshire and Coursey 1987; Coursey et al. 1987). In the present experiments, the markets for consumption goods were real and potentially binding from the first trial, and the WTA-WTP discrepancy was found to be stable over a series of such binding trials.

It should be stressed that previous research did not actually demonstrate that the discrepancy between buyers and sellers is eliminated in markets. Although the discrepancy between the final selling and buying prices in the sucrose octa-acetate experiment of Coursey et al. (1987) was not statistically significant, the ratio of median prices of sellers and buyers was still 2.6.[5] If the buyers and sellers had been allowed to trade according to their final bids, a total of nine advantageous exchanges would have occurred between the two groups, compared to the theoretical expectation of 16 trades (for details, see Knetsch and Sinden [1987]). This V/V^* ratio of .56 is quite similar to the ratios observed in experiments 1–4. In the study by Brookshire and Coursey (1987), the ratio of mean prices was indeed reduced by experience, from a high of 77 for initial hypothetical survey responses to 6.1 in the first potentially binding auction conducted in a laboratory. However, the ratio remained at 5.6 in the final auction.

III. Testing for Misrepresentation

As previously stated, subjects faced identical incentives in the induced-value and consumption goods phases of experiments 1–4. Therefore, it seems safe to attribute the difference in observed trading to the endowment effect. However, some readers of early drafts of this paper have suggested that because of the way market prices were determined, subjects might have felt that they had an incentive to misstate their true values in order to influence the price, and perhaps this incentive was perceived to be greater in the consumption goods markets. To eliminate this possible interpretation of the previous results, experiment 5 was carried out in a manner similar to the first four experiments, except that subjects were told that the price would be selected at random. As is well known, this is an incentive-compatible procedure for eliciting values (see Becker, DeGroot, and Marschak 1964).

Each participant received the following instructions (with appropriate alternative wording in the buyers' forms):

> After you have finished, one of the prices listed below will be selected at random and any exchanges will take place at that price. If you have indicated you will sell at this price you will receive this amount of money and will give up the mug; if you have indicated that you will keep the mug at this price

---

[5] The ratio of the mean selling and buying prices is 1.4 if all subjects are included. However, if one buyer and one seller with extreme valuations are excluded, the ratio is 1.9. These numbers were reported in an earlier version of Coursey et al. (1987).

> then no exchange will be made and you can take the mug home with you.
>
> . . . Your decision can have no effect on the price actually used because the price will be selected at random.

The experiment was conducted in a series of six tutorial groups of a business statistics class at Simon Fraser University. The use of small groups helped assure complete understanding of the instructions, and the exercises were conducted over the course of a single day to minimize opportunities for communication between participants. Each group was divided equally: half of the subjects were designated as sellers by random selection, and the other half became buyers. A total of 59 people took part.

Two induced-value markets for hypothetical payoffs and a subsequent third real exchange market for money and mugs were conducted with identical trading rules used in all three. All participants maintained the same role as either buyers or sellers for the three markets. As in experiments 1 and 2, the prices that individuals chose to buy or to sell were selected from possible prices ranging from $0.00 to $9.50 listed by increments of $0.50.

The results of this experiment were nearly identical to the earlier ones in which the actual exchanges were based on the market-clearing price. Even though possibly less motivating hypothetical values were used in the two induced-value markets, nearly all participants pursued a profit-maximizing selection of prices to buy or sell the assets. Fourteen exchanges at a price of $4.75 were expected in the first induced-value market on the basis of the randomly distributed values written on the forms. Thirteen trades at this price were indicated by the prices actually selected by the participants. The results of the second hypothetical induced-value market were equally convincing, with 16 of the 17 expected exchanges made at the expected price of $5.75. The procedures and incentives were apparently well understood by the participants.

Mugs, comparable to those used in other experiments, were distributed to the potential sellers after the induced-value markets were completed. A mug was also shown to all the potential buyers. The following form with instructions, nearly identical to the ones used in the induced-value markets, was then distributed (with the alternative wording for buyers in brackets):

> You now [do not] have, and own a mug which you can keep and take home. You also have the option of selling it and receiving [buying one to take home by paying] money for it.
>
> For each of the possible prices listed below, please indicate

whether you wish to: (1) Receive [pay] that amount of money and sell your [buy a] mug, or (2) Not sell your [buy a] mug at this price.

After you have finished, one of the prices listed below will be selected at random and any exchanges will take place at that price. If you have indicated you will sell [buy] at this price you will receive this amount of money [a mug] and will give up the mug [pay this amount of money]; if you have indicated that you will keep the [not buy a] mug at this price then no exchange will be made and you can take the mug home with you [do not pay anything].

Notice the following two things: (1) Your decision can have no effect on the price actually used because the price will be selected at random. (2) It is in your interest to indicate your true preferences at each of the possible prices listed below.

For each price indicate your decision by marking an X in the appropriate column.

	I Will Sell [Buy]	I Will Keep [Not Buy] the Mug
If the price is $0	_____	_____
If the price is $0.50	_____	_____
⋮		
If the price is $9.50	_____	_____

After the instructions were read and reviewed by the experimenter and questions were answered, participants completed the forms indicating either their lowest selling price or their highest buying price. A random price, from among the list from $0.00 to $9.50, was then drawn, and exchanges based on this price were completed.

The results again showed a large and significant endowment effect. Given the 29 potential buyers, 30 potential sellers, and the random distribution of the mugs, 14.5 exchanges would be expected if entitlements did not influence valuations. Instead, only six were indicated on the basis of the values actually selected by the potential buyers and sellers ($V/V^* = .41$). The median selling price of $5.75 was over twice the median buying price of $2.25, and the means were $5.78 and $2.21, respectively.

IV. Reluctance to Buy versus Reluctance to Sell

Exchanges of money and a good (or between two goods) offer the possibilities of four comparisons: a choice of gaining either the good or money, a choice of losing one or the other, buying (giving up money for the good), and selling (giving up the good for money) (Tversky and Kahneman, in press). The endowment effect results from a difference between the relative preferences for the good and money. The comparison of buying and selling to simple choices between gains permits an analysis of the discrepancy between WTA and WTP into two components: reluctance to sell (exchanging the good for money) and reluctance to buy (exchanging money for the good).

Experiments 6 and 7 were carried out to assess the weight of reluctance to buy and reluctance to sell in undertrading of a good similar to the goods used in the earlier experiments. The subjects in experiment 6 were 77 Simon Fraser students, randomly assigned to three groups. Members of one group, designated sellers, were given a coffee mug and were asked to indicate whether or not they would sell the mug at a series of prices ranging from $0.00 to $9.25. A group of buyers indicated whether they were willing to buy a mug at each of these prices. Finally, choosers were asked to choose, for each of the possible prices, between a mug and cash.

The results again reveal substantial undertrading: While 12.5 trades were expected between buyers and sellers, only three trades took place ($V/V^* = .24$). The median valuations were $7.12 for sellers, $3.12 for choosers, and $2.87 for buyers. The close similarity of results for buyers and choosers indicates that there was relatively little reluctance to pay for the mug.

Experiment 7 was carried out with 117 students at the University of British Columbia. It used an identical design except that price tags were left on the mugs. The results were consistent with those in experiment 6. Nineteen trades were expected on the basis of valuation equivalence, but only one was concluded on the basis of actual valuations ($V/V^* = .05$). The median valuations were $7.00 for sellers, $3.50 for choosers, and $2.00 for buyers.

It is worth noting that these results eliminate any form of income effect as an explanation of the discrepant valuations since the positions of sellers and choosers were strictly identical. The allocation of a particular mug to each seller evidently induced a sense of endowment that the choosers did not share: the median value of the mug to the sellers was more than double the value indicated by the choosers even though their choices were objectively the same. The results imply that the observed undertrading of consumption goods may be largely due to a reluctance to part with entitlements.

V. Bilateral Bargaining and the Coase Theorem

According to the Coase theorem, the allocation of resources to individuals who can bargain and transact at no cost should be independent of initial property rights. However, if the marginal rate of substitution between one good and another is affected by endowment, then the individual who is assigned the property right to a good will be more likely to retain it. A bilateral bargaining experiment (experiment 8) was carried out to test this implication of the endowment effect.

The subjects were 35 pairs of students in seven small tutorials at Simon Fraser University. The students were enrolled in either a beginning economics course or an English class. Each student was randomly paired with another student in the same tutorial group, with care taken to assure that students entering the tutorial together were not assigned as a pair. A game of Nim, a simple game easily explained, was played by each pair of participants. The winners of the game were each given a 400-gram Swiss chocolate bar and told it was theirs to keep.

An induced-value bargaining session was then conducted. The member of each pair who did not win the Nim game, and therefore did not receive the chocolate bar, was given a ticket and an instruction sheet that indicated that the ticket was worth $3.00 because it could be redeemed for that sum. The ticket owners were also told that they could sell the ticket to their partner if mutually agreeable terms could be reached. The partners (the chocolate bar owners) received instructions indicating that they could receive $5.00 for the ticket if they could successfully buy it from the owner. Thus there was a $2.00 surplus available to any pair completing a trade.

The pairs were then given an unlimited amount of time to bargain. Subjects were told that both credit and change were available from the experimenter. Results of the bargaining sessions were recorded on their instruction sheets.

Of the 35 pairs of participants, 29 agreed to an exchange ($V/V^* =$ .83). The average price paid for the 29 tickets was $4.09, with 12 of the exchange prices being exactly $4.00. Payments of the redemption values of the tickets were made as soon as the exchanges were completed. These payments were made in single dollar bills to facilitate trading in the subsequent bargaining session. After the ticket exchanges were completed, owners of the chocolate bars were told that they could sell them to their partners if a mutually agreeable price could be determined. The procedures used for the tickets were once again applied to these bargaining sessions.

An important effect of the preliminary induced-value ticket bargains was to provide the ticket owners with some cash. The average gain to the ticket owners (including the six who did not sell their tickets) was $3.90. The average gain to their partners (the chocolate bar owners) was only $0.76. Thus the potential chocolate bar buyers were endowed with an average of \$3.14 more than the owners, creating a small income effect toward the buyers. Also, to the extent that a windfall gain such as this is spent more casually by subjects than other money (for evidence on such a "house money effect," see Thaler and Johnson [1990]), trading of chocolate bars should be facilitated.

Results of the chocolate bar bargains once again suggest reluctance to trade. Rather than the 17.5 trades expected from the random allocations, only seven were observed ($V/V^* = .4$). The average price paid in those exchanges that did occur was \$2.69 (the actual prices were \$6.00, \$3.10, \$3.00, \$2.75, \$2.00, \$1.00, and \$1.00). If the six pairs of subjects who did not successfully complete bargains in the first stage are omitted from the sample on the grounds that they did not understand the task or procedures, then six trades are observed where 14.5 would be expected ($V/V^* = .414$). Similarly, if two more pairs are dropped because the prices at which they exchanged tickets were outside the range \$3.00–\$5.00, then the number of trades falls to four, and V/V^* falls to .296. (No significant differences between the students in the English and economics classes were observed.)[6]

To be sure that the chocolate bars were valued by the subjects and that these valuations would vary enough to yield mutually beneficial trades, the same chocolate bars were distributed to half the members of another class at Simon Fraser. Those who received chocolate bars were asked the minimum price they would accept to sell their bar, while those without the bars were asked the maximum price they would pay to acquire a bar. The valuations of the bars varied from \$0.50 to \$8.00. The average value ascribed by sellers was \$3.98, while the buyers' average valuation was \$1.25. (The median values were \$3.50 and \$1.25.)

VI. The Endowment Effect in Choices between Goods

The previous experiments documented undertrading in exchanges of money and consumption goods. A separate experiment (Knetsch 1989) establishes the same effect in exchanges between two goods. Participants in three classes were offered a choice between the same two goods. All students in one class were given a coffee mug at the beginning of the session as compensation for completing a short questionnaire. At the completion of the task, the experimenters showed the students a bar of Swiss chocolate that they could immediately receive in exchange for the mug. The students in another class were offered an opportunity to make the opposite exchange after first being given the chocolate bar. The students in a third class were simply offered a choice, at the beginning of the session, between a chocolate bar and a mug. The proportion of students selecting the mug was 89 percent in the class originally endowed with mugs ($N = 76$), 56 percent in the class offered a choice ($N = 55$), and only 10 percent in the class originally endowed with chocolate bars ($N = 87$). For most participants a mug was more valuable than the chocolate when the mug had to be given up but less valuable when the chocolate had to be given up. This experiment confirms that undertrading can occur even when income effects are ruled out. It also demonstrates an endowment effect for a good that was distributed to everyone in the class and therefore did not have the appeal of a prize or trophy.

---

[6] We conducted two similar bargaining experiments that yielded comparable results. Twenty-six pairs of subjects negotiated the sale of mugs and then envelopes containing an uncertain amount of money. Buyers had not been given any cash endowment. These sessions yielded six and five trades, respectively, where 13 would be expected. Also, some induced-value bilateral negotiation sessions were conducted in which only \$0.50 of surplus was available (the seller's valuation was \$1.50 and the buyer's was \$2.00). Nevertheless, 21 of a possible 26 trades were completed.

VII. Discussion

The evidence presented in this paper supports what may be called an instant endowment effect: the value that an individual assigns to such objects as mugs, pens, binoculars, and chocolate bars appears to increase substantially as soon as that individual is given the object.[7] The apparently instantaneous nature of the reference point shift and consequent value change induced by giving a person possession of a good goes beyond previous discussions of the endowment effect, which focused on goods that have been in the individual's possession for some time. While long-term endowment effects could be explained by sentimental attachment or by an improved technology of consumption in the Stigler-Becker (1977) sense, the differences in preference or taste demonstrated by more than 700 participants in the experiments reported in this paper cannot be explained in this fashion.

The endowment effect is one explanation for the systematic differences between buying and selling prices that have been observed so often in past work. One of the objectives of this study was to examine an alternative explanation for this buying-selling discrepancy, namely that it reflects a general bargaining strategy (Knez and Smith 1987) that would be eliminated by experience in the market (Brookshire and Coursey 1987; Coursey et al. 1987). Our results do not support this alternative view. The trading institution used in experiments 1–7 encouraged participants to be price takers (especially in experiment 5), and the rules provided no incentive to conceal true preferences. Furthermore, the results of the induced-value markets indicate that the subjects understood the demand-revealing nature of the questions they were asked and acted accordingly. Substantial undertrading was nevertheless observed in markets for consumption goods. As for learning and market discipline, there was no indication that buying and selling prices converged over repeated market trials, though full feedback was provided at the end of each trial. The undertrading observed in these experiments appears to reflect a true difference in preferences between the potential buyers and sellers. The robustness of this result reduces the risk that the outcome is produced by an experimental artifact. In short, the present findings indicate that the endowment effect can persist in genuine market settings.

---

[7] The impression gained from informal pilot experiments is that the act of giving the participant physical possession of the good results in a more consistent endowment effect. Assigning subjects a chance to receive a good, or a property right to a good to be received at a later time, seemed to produce weaker effects.

The contrast between the induced-value markets and the consumption goods markets lends support to Heiner's (1985) conjecture that the results of induced-value experiments may not generalize to all market settings. The defining characteristic of the induced-value markets is that the values of the tokens are unequivocally defined by the amount the experimenter will pay for them. Loss aversion is irrelevant with such objects because transactions are evaluated simply on the basis of net gain or loss. (If someone is offered $6.00 for a $5.00 bill, there is no sense of loss associated with the trade.) Some markets may share this feature of induced-value markets, especially when the conditions of pure arbitrage are approached. However, the computation of net gain and loss is not possible in other situations, for example, in markets in which risky prospects are traded for cash or in markets in which people sell goods that they also value for their use. In these conditions, the cancellation of the loss of the object against the dollars received is not possible because the good and money are not strictly commensurate. The valuation ambiguity produced by this lack of commensurability is necessary, although not sufficient, for both loss aversion and a buying-selling discrepancy.

The results of the experimental demonstrations of the endowment effect have direct implications for economic theory and economic predictions. Contrary to the assumptions of standard economic theory that preferences are independent of entitlements,[8] the evidence presented here indicates that people's preferences depend on their reference positions. Consequently, preference orderings are not defined independently of endowments: good A may be preferred to B when A is part of an original endowment, but the reverse may be true when initial reference positions are changed. Indifference curves will have a kink at the endowment or reference point (see Tversky and Kahneman, in press), and an indifference curve tracing acceptable trades in one direction may even cross another indifference curve that plots the acceptable exchanges in the opposite direction (Knetsch 1989).

---

[8] Although ownership can affect taste in the manner suggested by Stigler and Becker (1977), in the absence of income effects, it is traditional to assume that the indifference curves in an Edgeworth box diagram do not depend on the location of the endowment point.

The existence of endowment effects reduces the gains from trade. In comparison with a world in which preferences are independent of endowment, the existence of loss aversion produces an inertia in the economy because potential traders are more reluctant to trade than is conventionally assumed. This is not to say that Pareto-optimal trades will not take place. Rather, there are simply fewer mutually advantageous exchanges possible, and so the volume of trade is lower than it otherwise would be.

To assess the practical significance of the endowment effect, it is important to consider first some necessary conditions for the effect to be observed. Experiments 6 and 7 suggest that the endowment effect is primarily a problem for sellers; we observed little reluctance to buy but much reluctance to sell. Furthermore, not all sellers are afflicted by an endowment effect. The effect did not appear in the markets for money tokens, and there is no reason in general to expect reluctance to resell goods that are held especially for that purpose. An owner will not be reluctant to sell an item at a given price if a perfect substitute is readily available at a lower price. This reasoning suggests that endowment effects will almost certainly occur when owners are faced with an opportunity to sell an item purchased for use that is not easily replaceable. Examples might include tickets to a sold-out event, hunting licenses in limited supply (Bishop and Heberlein 1979), works of art, or a pleasant view.

While the conditions necessary for an endowment effect to be observed may appear to limit its applicability in economic settings, in fact these conditions are very often satisfied, and especially so in the bargaining contexts to which the Coase theorem is applied. For example, tickets to Wimbledon are allocated by means of a lottery. A standard Coasean analysis would imply that in the presence of an efficient ticket brokerage market, winners of the lottery would be no more likely to attend the matches than other tennis fans who had won a similar cash prize in an unrelated lottery. In contrast, the experimental results presented in this paper predict that many winners of Wimbledon tickets will attend the event, turning down opportunities to sell their tickets that exceed their reservation price for buying them.

Endowment effects can also be observed for firms and other organizations. Endowment effects are predicted for property rights acquired by historic accident or fortuitous circumstances, such as government licenses, landing rights, or transferable pollution permits. Owing to endowment effects, firms will be reluctant to divest themselves of divisions, plants, and product lines even though they would never consider buying the same assets; indeed, stock prices often rise when firms do give them up. Again, the prediction is not an absence of trade, just a reduction in the volume of trade.

Isolating the influence of endowment effects from those of transaction costs as causes of low trading volumes is, of course, difficult in actual market settings. Demonstrations of endowment effects are most persuasive where transaction costs are very small. By design, this was the case in the experimental markets, where the efficiency of the induced-value markets demonstrated the minimal effect of transaction costs, or other impediments, on exchange decisions, leaving the great reluctance to trade mugs and other goods to be attributable to endowment effects.

Endowment effects are not limited to cases involving physical goods or to legal entitlements. The reference position of individuals and firms often includes terms of previous transactions or expectations of continuation of present, often informal, arrangements. There is clear evidence of dramatically asymmetric reactions to improvements and deteriorations of these terms and a willingness to make sacrifices to avoid unfair treatment (Kahneman, Knetsch, and Thaler 1986). The reluctance to sell at a loss, owing to a perceived entitlement to a formerly prevailing price, can explain two observations of apparent undertrading. The first pertains to housing markets. It is often observed that when housing prices fall, volume also falls. When house prices are falling, houses remain on the market longer than when prices are rising. Similarly, the volume for stocks that have declined in price is lower than the volume for stocks that have increased in value (Shefrin and Statman 1985; Ferris, Haugen, and Makhija 1988), although tax considerations would lead to the opposite prediction.

Another manifestation of loss aversion in the context of multiattribute negotiations is what might be termed "concession aversion": a reluctance to accept a loss on any dimension of an agreement. A straightforward and common instance of this is the downward stickiness of wages. A somewhat more subtle implication of concession aversion is that it can produce inefficient contract terms owing to historic precedents. Old firms may have more inefficient arrangements than new ones because new companies can negotiate without the reference positions created by prior agreements. Some airlines, for example, are required to carry three pilots on some planes while others—newer ones—operate with two.

Loss aversion implies a marked asymmetry in the treatment of losses and forgone gains, which plays an essential role in judgments of fairness (Kahneman et al. 1986). Accordingly, disputes in which concessions are viewed as losses are often much less tractable than disputes in which concessions involve forgone gains. Court decisions recognize the asymmetry of losses and forgone gains by favoring possessors of goods over other claimants, by limiting recovery of lost profits relative to compensation for actual expenditures, and by failing to enforce gratuitous promises that are coded as forgone gains to the injured party (Cohen and Knetsch 1989).

To conclude, the evidence reported here offers no support for the contention that observations of loss aversion and the consequential evaluation disparities are artifacts; nor should they be interpreted as mistakes likely to be eliminated by experience, training, or "market discipline." Instead, the findings support an alternative view of endowment effects and loss aversion as fundamental characteristics of preferences.

References

Banford, Nancy D.; Knetsch, Jack L.; and Mauser, Gary A. "Feasibility Judgements and Alternative Measures of Benefits and Costs." *J. Bus. Admin.* 11, nos. 1, 2 (1979): 25–35.

Becker, Gordon M.; DeGroot, Morris H.; and Marschak, Jacob. "Measuring Utility by a Single-Response Sequential Method." *Behavioral Sci.* 9 (July 1964): 226–32.

Bishop, Richard C., and Heberlein, Thomas A. "Measuring Values of Extramarket Goods: Are Indirect Measures Biased?" *American J. Agricultural Econ.* 61 (December 1979): 926–30.

Boyce, Rebecca R.; Brown, Thomas C.; McClelland, Gary D.; Peterson, George L.; and Schulze, William D. "An Experimental Examination of Intrinsic Environmental Values." Working paper. Boulder: Univ. Colorado, 1990.

Brookshire, David S., and Coursey, Don L. "Measuring the Value of a Public Good: An Empirical Comparison of Elicitation Procedures." *A.E.R.* 77 (September 1987): 554–66.

Brookshire, David S.; Randall, Alan; and Stoll, John R. "Valuing Increments and Decrements in Natural Resource Service Flows." *American J. Agricultural Econ.* 62 (August 1980): 478–88.

Cohen, David, and Knetsch, Jack L. "Judicial Choice and Disparities between Measures of Economic Values." Working paper. Burnaby, B.C.: Simon Fraser Univ., 1989.

Coursey, Don L.; Hovis, John L.; and Schulze, William D. "The Disparity between Willingness to Accept and Willingness to Pay Measures of Value." *Q.J.E.* 102 (August 1987): 679–90.

Ferris, Stephen P.; Haugen, Robert A.; and Makhija, Anil K. "Predicting Contemporary Volume with Historic Volume at Differential Price Levels: Evidence Supporting the Disposition Effect." *J. Finance* 43 (July 1988): 677–97.

Freeman, A. Myrick. *The Benefits of Environmental Improvement.* Washington: Resources for the Future, 1979.

Hammack, Judd. and Brown, Gardner Mallard, Jr. *Waterfowl and Wetlands: Toward Bio-economic Analysis.* Baltimore: Johns Hopkins Press (for Resources for the Future), 1974.

Heberlein, Thomas A., and Bishop, Richard C. "Assessing the Validity of Contingent Valuation: Three Field Experiments." Paper presented at the International Conference on Man's Role in Changing the Global Environment, Italy, 1985.

Heiner, Ronald A. "Experimental Economics: Comment." *A.E.R.* 75 (March 1985): 260–63.

Kahneman, Daniel; Knetsch, Jack L.; and Thaler, Richard. "Fairness as a Constraint on Profit Seeking: Entitlements in the Market." *A.E.R.* 76 (September 1986): 728–41.

Kahneman, Daniel, and Tversky, Amos. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (March 1979): 263–91.

Knetsch, Jack L. "The Endowment Effect and Evidence of Nonreversible Indifference Curves." *A.E.R.* 79 (December 1989): 1277–84.

Knetsch, Jack L., and Sinden, J. A. "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value." *Q.J.E.* 99 (August 1984): 507–21.

———. "The Persistence of Evaluation Disparities." *Q.J.E.* 102 (August 1987): 691–95.

Knez, Marc, and Smith, Vernon L. "Hypothetical Valuations and Preference Reversals in the Context of Asset Trading." In *Laboratory Experiments in Economics: Six Points of View,* edited by Alvin E. Roth. Cambridge: Cambridge Univ. Press, 1987.

Knez, Peter; Smith, Vernon L.; and Williams, Arlington W. "Individual Rationality, Market Rationality, and Value Estimation." *A.E.R. Papers and Proc.* 75 (May 1985): 397–402.

Plott, Charles R. "Industrial Organization Theory and Experimental Economics." *J. Econ. Literature* 20 (December 1982): 1485–1527.

Rowe, Robert D.; d'Arge, Ralph C.; and Brookshire, David S. "An Experiment on the Economic Value of Visibility." *J. Environmental Econ. and Management* 7 (March 1980): 1–19.

Shefrin, Hersh, and Statman, Meir. "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence." *J. Finance* 40 (July 1985): 777–90.

Sinclair, William F. *The Economic and Social Impact of Kemano II Hydroelectric Project on British Columbia's Fisheries Resources.* Vancouver: Dept. Fisheries and Oceans, 1978.

Smith, Vernon L. "Experimental Economics: Induced Value Theory." *A.E.R. Papers and Proc.* 66 (May 1976): 274–79.

———. "Macroeconomic Systems as an Experimental Science." *A.E.R.* 72 (December 1982): 923–55.

Stigler, George J., and Becker, Gary S. "De Gustibus Non Est Disputandum." *A.E.R.* 67 (March 1977): 76–90.

Thaler, Richard. "Toward a Positive Theory of Consumer Choice." *J. Econ. Behavior and Organization* 1 (March 1980): 39–60.

Thaler, Richard, and Johnson, Eric J. "Gambling with the House Money and Trying to Break Even: The Effects of Prior Outcomes on Risky Choice." *Management Sci.* 36 (June 1990).

Tversky, Amos, and Kahneman, Daniel. "Reference Theory of Choice and Exchange." *Q.J.E.* (in press).

Willig, Robert D. "Consumer's Surplus without Apology." *A.E.R.* 66 (September 1976): 589–97.

ATTACHMENT C

Journal of Economic Perspectives—Volume 5, Number 1—Winter 1991—Pages 193–206

Anomalies
The Endowment Effect, Loss Aversion, and Status Quo Bias

Daniel Kahneman, Jack L. Knetsch,
and Richard H. Thaler

Economics can be distinguished from other social sciences by the belief that most (all?) behavior can be explained by assuming that agents have stable, well-defined preferences and make rational choices consistent with those preferences in markets that (eventually) clear. An empirical result qualifies as an anomaly if it is difficult to "rationalize," or if implausible assumptions are necessary to explain it within the paradigm. This column presents a series of such anomalies. Readers are invited to suggest topics for future columns by sending a note with some reference to (or better yet copies of) the relevant research. Comments on anomalies printed here are also welcome. The address is: Richard Thaler, c/o Journal of Economic Perspectives, Johnson Graduate School of Management, Malott Hall, Cornell University, Ithaca, NY 14853.

After this issue, the "Anomalies" column will no longer appear in every issue and instead will appear occasionally, when a pressing anomaly crosses Dick Thaler's desk. However, suggestions for new columns and comments on old ones are still welcome. Thaler would like to quash one rumor before it gets started, namely that he is cutting back because he has run out of anomalies. *Au contraire*, it is the dilemma of choosing which juicy anomaly to discuss that takes so much time.

▪ *Daniel Kahneman is Professor of Psychology, University of California, Berkeley, California. Jack L. Knetsch is Professor of Economics and Natural Resources Management, Simon Fraser University, Burnaby, British Columbia, Canada. Richard H. Thaler is Henrietta Johnson Louis Professor of Economics, Johnson School of Management, Cornell University, Ithaca, New York.*

Introduction

A wine-loving economist we know purchased some nice Bordeaux wines years ago at low prices. The wines have greatly appreciated in value, so that a bottle that cost only $10 when purchased would now fetch $200 at auction. This economist now drinks some of this wine occasionally, but would neither be willing to sell the wine at the auction price nor buy an additional bottle at that price.

Thaler (1980) called this pattern—the fact that people often demand much more to give up an object than they would be willing to pay to acquire it—the *endowment effect*. The example also illustrates what Samuelson and Zeckhauser (1988) call a *status quo bias*, a preference for the current state that biases the economist against both buying *and* selling his wine. These anomalies are a manifestation of an asymmetry of value that Kahneman and Tversky (1984) call *loss aversion*—the disutility of giving up an object is greater that the utility associated with acquiring it. This column documents the evidence supporting endowment effects and status quo biases, and discusses their relation to loss aversion.

The Endowment Effect

An early laboratory demonstration of the endowment effect was offered by Knetsch and Sinden (1984). The participants in this study were endowed with either a lottery ticket or with $2.00. Some time later, each subject was offered an opportunity to trade the lottery ticket for the money, or vice versa. Very few subjects chose to switch. Those who were given lottery tickets seemed to like them better than those who were given money.

This demonstration and other similar ones (Knetsch, 1989), while striking, did not settle the matter. Some economists felt that the behavior would disappear if subjects were exposed to a market environment with ample learning opportunities. For example, Knez, Smith and Williams (1985) argued that the discrepancy between buying and selling prices might be produced by the thoughtless application of normally sensible bargaining habits, namely understating one's true willingness to pay (WTP) and overstating the minimum acceptable price at which one would sell (willingness to accept or WTA). Coursey, Hovis, and Schultze (1987) reported that the discrepancy between WTP and WTA diminished with experience in a market setting (although it was probably not eliminated, see Knetsch and Sinden, 1987). To clarify the issue, Kahneman, Knetsch, and Thaler (1990) ran a new series of experiments to determine whether the endowment effect survives when subjects face market discipline and have a chance to learn. We will report just two experiments from that series.

In the first experiment, students in an advanced undergraduate economics class at Cornell University participated in a series of markets. The objects traded in the first three markets were 'induced value tokens.' In such markets all subjects are told how much a token is worth to them, with the amounts varying across subjects. Half the subjects were made owners of tokens, the other half were not. In this way, supply and demand curves for tokens are created. Subjects alternated between the buyer and seller role in the three successive markets, and were assigned a different individual redemption value in each trial. Experimenters collected the forms from all participants after each market period, and immediately calculated and announced the market-clearing price and the number of trades. Three buyers and three sellers were selected at random after each of the induced markets and were paid off according to the preferences stated on their forms and the market clearing price for that period.

These markets contained no grist for the anomaly mill. On each trial, the market clearing price was exactly equal to the intersection of the induced supply and demand curves, and the volume of trade was within one unit of the predicted quantity. These results demonstrate that the subjects understood the task, and that the market mechanism used did not impose high transactions costs.

Immediately after the three induced value markets, subjects on alternating seats were given Cornell coffee mugs, which sell for $6.00 each at the bookstore. The experimenter asked all participants to examine a mug, either their own or their neighbor's. The experimenter then informed the subjects that four markets for mugs would be conducted using the same procedures as the prior induced markets with two exceptions: (1) One of the four market trials would subsequently be selected at random and only the trades made on this trial would be executed. (2) On the binding market trial, *all* trades would be implemented, unlike the subset implemented in the induced value markets. The initial assignment of buyer and seller roles was maintained for all four trading periods. The clearing price and the number of trades were announced after each period. The market that "counted" was indicated after the fourth period, and transactions were executed immediately—all sellers who had indicated that they would give up their mug at the market clearing price exchanged their mugs for cash, and successful buyers paid this same price and received their mug. This design was used to permit learning to take place over successive trials and yet make each trial potentially binding. The same procedure was then followed for four more successive markets using boxed ball-point pens with a visible bookstore price tag of $3.98, which were distributed to the subjects who had been buyers in the mug markets.

What does economic theory predict will happen in these markets for mugs and pens? Since transactions costs have been shown to be insignificant in the induced value markets, and income effects are trivial, a clear prediction is available: When the market clears, the objects will be owned by those subjects who value them most. Call the half of the subjects who like mugs the most "mug lovers" and the half who like mugs least "mug haters." Then, since the mugs were assigned at random, on average half of the mug lovers will be given a mug, and half will not. This implies that in the market, half of the mugs should trade, with mug haters selling to mug lovers.

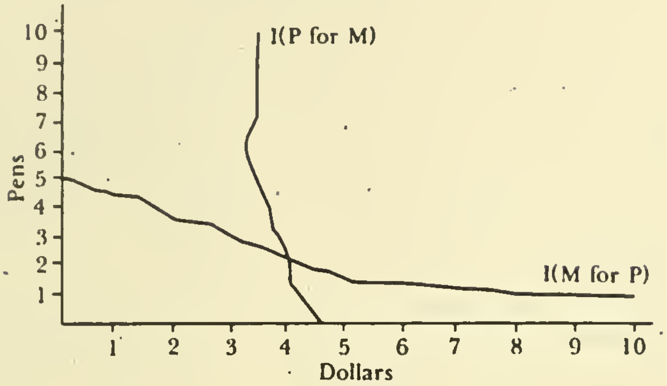
The 50 percent predicted volume of trade did not materialize. There were 22 mugs and pens distributed so the predicted number of trades was 11. In the four mug markets the number of trades was 4, 1, 2, and 2 respectively. In the pen markets the number of trades was either 4 or 5. In neither market was there any evidence of a trend over the four trials. The reason for the low volume of trade is revealed by the reservation prices of buyers and sellers. For mugs, the median owner was unwilling to sell for less than $5.25, while the median buyer was unwilling to pay more than $2.25-$2.75. The market price varied between $4.25 and $4.75. In the market for pens the ratio of selling to buying prices was also about 2. The experiment was replicated several times, always with similar results: median selling prices are about twice median buying prices and volume is less than half of that expected.

Another experiment from this series allows us to investigate whether the low volume of trading is produced by a reluctance to buy or a reluctance to sell. In this experiment, 77 students at Simon Fraser University were randomly assigned to three conditions. One group, the Sellers, were given SFU coffee mugs and were asked whether they would be willing to sell the mugs at each of a series of prices ranging from $0.25 to $9.25. A second group of Buyers were asked whether they would be willing to buy a mug at the same set of prices. The third group, called Choosers, were not given a mug but were asked to choose, for each of the prices, between receiving a mug or that amount of money.

Notice that the Sellers and the Choosers are in objectively identical situations, deciding at each price between the mug and that amount of money. Nevertheless, the Choosers behaved more like Buyers than like Sellers. The median reservation prices were: Sellers, $7.12; Choosers, $3.12; Buyers, $2.87. This suggests that the low volume of trade is produced mainly by owner's reluctance to part with their endowment, rather than by buyers' unwillingness to part with their cash. This experiment also eliminates the trivial income effect present in the first experiment, since the Sellers and Choosers are in the same economic situation.

One of the first lessons in microeconomics is that two indifference curves can never intersect. This result depends on the implicit assumption that indifference curves are reversible. That is, if an individual owns x and is indifferent between keeping it and trading it for y, then when owning y the individual should be indifferent about trading it for x. If loss aversion is present, however, this reversibility will no longer hold. Knetsch (1990) has demonstrated this point experimentally. One group of subjects received 5 medium priced ball point pens, while another group of subjects received $4.50. They were then made a series of offers which they could accept or reject. The offers were designed to identify an indifference curve. For example, someone

Figure 1
Crossing indifference curves



who had been given the pens would be asked if she would give up one of the pens for a dollar. One of the accepted offers (including the original endowment) was selected at random at the end of the experiment to determine the subject's payment. By plotting the line between accepted and rejected offers, Knetsch was able to infer an indifference curve for each subject. Then he plotted the average indifference curve for each of the two groups (those who started with pens and those who started with money). These plots are shown in Figure 1. The curves are quite different: the pens were worth more money to those subjects who started with pens than to those who started with money. As a result, the curves intersect.[1]

What produces these "instant endowment effects"? Do subjects who receive a gift actually value it more than others who do not receive it? A recent study by Loewenstein and Kahneman (1991) investigated this issue. Half the students in a class ($N = 63$) were given pens, the others were given a token redeemable for an unspecified gift. All participants were then asked to rank the attractiveness of six gifts under consideration as prizes in subsequent experiments. Finally, all the subjects were then given a choice between a pen and two chocolate bars. As in previous experiments, there was a pronounced endowment effect. The pen was preferred by 56 percent of those endowed with it, but only 24 percent of the other subjects chose a pen. However, when making the attractiveness ratings, the subjects endowed with pens did not rate them as more attractive. This suggests that the main effect of endowment is not to enhance the appeal of the good one owns, only the pain of giving it up.

Status Quo Bias

One implication of loss aversion is that individuals have a strong tendency to remain at the status quo, because the disadvantages of leaving it loom larger than advantages. Samuelson and Zeckhauser (1988) have demonstrated this effect, which they term the *status quo bias*. In one experiment, some subjects were given a hypothetical choice task, such as the following, in a 'neutral' version in which no status quo is defined: "You are a serious reader of the financial pages but until recently have had few funds to invest. That is when you inherited a large sum of money from your great-uncle. You are considering different portfolios. Your choices are to invest in: a moderate-risk company, a high risk company, treasury bills, municipal bonds."

---

[1] These curves were obtained from different individuals. Because subjects were randomly assigned to the two endowment groups, however, it is reasonable to attribute crossing indifference curves to the representative individual.

Other subjects were presented with the same problem but with one of the options designated as the status quo. In this case, after the same opening sentence the passage continues: "...That is when you inherited a portfolio of cash and securities from your great-uncle. A significant portion of this portfolio is invested in a moderate risk company...(The tax and broker commission consequences of any change are insignificant.)"

Many different scenarios were investigated, all using the same basic experimental design. Aggregating across all the different questions, Samuelson and Zeckhauser are then able to estimate the probability that an option will be selected when it is the status quo or when it is competing as an alternative to the status quo as a function of how often it is selected in the neutral setting. Their results implied that an alternative became significantly more popular when it was designated as the status quo. Also, the advantage of the status quo increases with the number of alternatives.

A test of status quo bias in a field setting was performed by Hartman, Doane, and Woo (forthcoming) using a survey of California electric power consumers. The consumers were asked about their preferences regarding service reliability and rates. They were told that their answers would help determine company policy in the future. The respondents fell into two groups, one with much more reliable service than the other. Each group was asked to state a preference among six combinations of service reliabilities and rates, with one of the combinations designated as the status quo. The results demonstrated a pronounced status quo bias. In the high reliability group, 60.2 percent selected their status quo as their first choice, while only 5.7 percent expressed a preference for the low reliability option currently being experienced by the other group, though it came with a 30 percent reduction in rates. The low reliability group, however, quite liked their status quo, 58.3 percent of them ranking it first. Only 5.8 percent of this group selected the high reliability option at a proposed 30 percent increase in rates.[2]

---

[2] Differences in income and electricity consumption between the two groups were minor and did not appear to significantly influence the results. Could the results be explained by either learning or habituation? That is, might the low reliability group have learned to cope with frequent outages, or found out that candlelight dinners are romantic? This cannot be ruled out, but it should be stressed that no similar explanation can be used for the mug experiments or the surveys conducted by Samuelson and Zeckhauser, so at least some of the effects observed are attributable to a pure status quo bias.

A large-scale experiment on status quo bias is now being conducted (inadvertently) by the states of New Jersey and Pennsylvania. Both states now offer a choice between two types of automobile insurance: a cheaper policy that restricts the right to sue, and a more expensive one that maintains the unrestricted right. Motorists in New Jersey are offered the cheaper policy as the default option, with an opportunity to acquire an unrestricted right to sue at a higher price. Since this option was made available in 1988, 83 percent of the drivers have elected the default option. In Pennsylvania's 1990 law, however, the default option is the expensive policy, with an opportunity to opt for the cheaper kind. The potential effect of this legislative framing manipulation was studied by Hershey, Johnson, Meszaros, and Robinson (1990). They asked two groups to choose between alternative policies. One group was presented with the New Jersey plan while the other was presented with the Pennsylvania plan. Of those subjects offered the New Jersey plan, only 23 percent elected to buy the right to sue whereas 53 percent of the subjects offered the Pennsylvania plan retained that right. On the basis of this research, the authors predict that more Pennsylvanians will elect the right to sue than New Jerseyans. Time will tell.

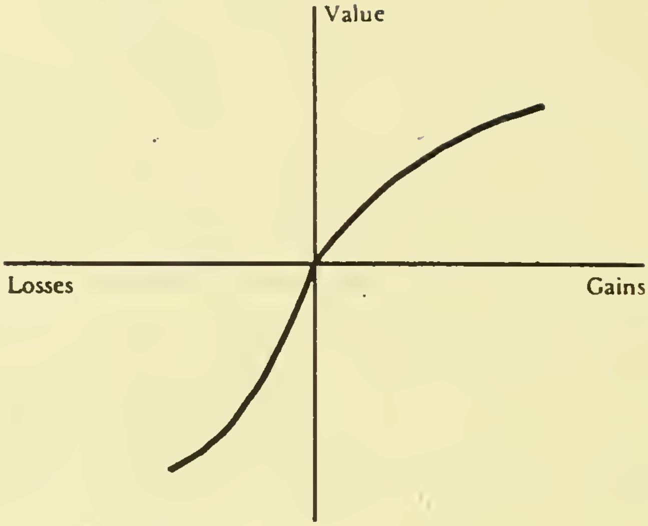
One final example of a presumed status quo bias comes courtesy of the *JEP* staff. Among Carl Shapiro's comments on this column was this gem: "You may be interested to know that when the AEA was considering letting members elect to drop one of the three Association journals and get a credit, prominent economists involved in that decision clearly took the view that fewer members would choose to drop a journal if the default was presented as all three journals (rather than the default being 2 journals with an extra charge for getting all three). We're talking economists here."

Loss Aversion

These observations, and many others, can be explained by a notion of loss aversion. A central conclusion of the study of risky choice has been that such choices are best explained by assuming that the significant carriers of utility are not states of wealth or welfare, but changes relative to a neutral reference point. Another central result is that changes that make things worse (losses) loom larger than improvements or gains. The choice data imply an abrupt change of the slope of the value function at the origin. The existing evidence suggests that the ratio of the slopes of the value function in two domains, for small or moderate gains and losses of money, is about 2:1 (Tversky and Kahneman, 1991). A schematic value function is shown in Figure 2.

The natural extension of this idea to riskless choice is that the attributes of options in trades and other transactions are also evaluated as gains and losses relative to a neutral reference point. The approach is illustrated in Figure 3. Decision makers have a choice between state *A*, where they have more of good

Figure 2
A typical value function



Y and less of good *X*, and state *D*, where they have more of good *X* and less of good *Y*. Four different reference points are indicated in the Figure. The individual faces a positive choice between two gains if the reference point is *C*, a negative choice between two losses if the reference point is *B*, and two different exchanges if the references are *A* or *D*, respectively. For example, if good *Y* is a mug and good *X* is money, the reference points for the sellers and the choosers in the mugs experiment are *A* and *C*. Loss aversion implies that the difference between the states of having a mug and not having one is larger from *A* than from *C*, which explains the different monetary values that subjects attach to the mug in these conditions.[3] For a formal treatment that generalizes consumer theory by introducing the notions of reference and loss aversion, see Tversky and Kahneman (1991).

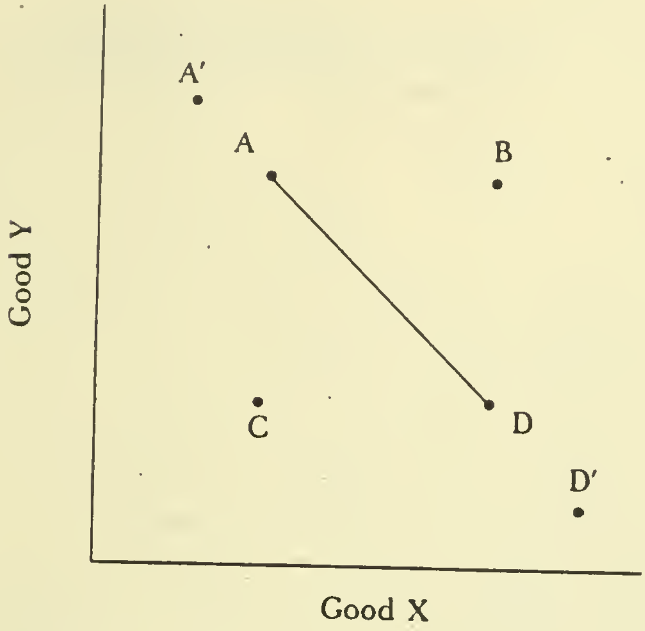
In general, a given difference between two options will have greater impact if it is viewed as a difference between two disadvantages than if it is viewed as a difference between two advantages. The status quo bias is a natural consequence of this asymmetry: the disadvantages of a change loom larger than its advantages. However, the differential weighting of advantages and disadvantages can be demonstrated even when the retention of the status quo is not an option. For an example, consider the following question (from Tversky and Kahneman, 1991):

> Imagine that as part of your professional training you were assigned to a part-time job. The training is now ending and you must look for employ-

---

[3] Loss aversion does not affect all transactions. In a normal commercial transaction, the seller does not suffer a loss when trading a good. Furthermore, the evidence indicates that buyers do not value the money spent on normal purchases as a loss, so long as the price of the good is not thought to be unusually high. Loss aversion is expected to primarily affect owners of goods that had been bought for use rather than for eventual resale.

Figure 3
Multiple reference points for the choice between A and D



ment. You consider two possibilities. They are like your training job in most respects except for the amount of social contact and the convenience of commuting to and from work. To compare the two jobs to each other and to the present one you have made up the following table:

Job	Contact with others	Commute Time
Present job	isolated for long stretches	10 min.
Job *A*	limited contact with others	20 min.
Job *D*	moderately sociable	60 min.

The options *A* and *D* are evaluated from a reference job which is better on commute time and worse on personal contact (a point like *A'* in Figure 3). Another version of the problem presented the same options, but the reference job involved "much pleasant social interaction and 80 minutes of daily commuting time," which corresponds to the point *D'*. The proportion of subjects choosing job *A* was 70 percent in the first version, 33 percent in the second. Subjects are more sensitive to the dimension in which they are losing relative to their reference point.

Some asymmetries between buying and selling prices are much too large to be explained by garden-variety loss aversion. For example, Thaler (1980) told subjects that they had been exposed to a rare fatal disease and that they now face a .001 chance of painless death within two weeks. They must decide how much they would be willing to pay for a vaccine, to be purchased immediately. The same subjects were also asked for the compensation they would demand to participate in a medical experiment in which they face a .001 chance of a quick and painless death. For most subjects the two prices differed by more than an order of magnitude.

A study by Viscusi, Magat and Huber (1987) documented a similar effect in a more realistic setting. Their respondents were recruited at a shopping mall and hardware store. The respondents were shown a can of fictitious insecticide, and were asked to examine it for their use. The current price of the can was said to be $10. Respondents were informed that all insecticide can cause injuries if misused, including inhalation and skin poisoning (in households with young children, child poisoning replaced skin poisoning). The current risk level was said to be 15 injuries of each type per 10,000 bottles sold. Respondents were asked to state their willingness-to-pay (WTP) to eliminate or reduce the risks. In households without children, the mean WTP to eliminate both risks was $3.78. The respondents were also asked to state the price reduction they would require to accept an increase of 1/10,000 in each of the two risks. The results were dramatic: 77 percent of respondents in this condition said they would refuse to buy the product at any positive price.

The striking difference between WTA and WTP in these studies probably reflects the large difference in the responsibility costs associated with voluntary assumption of additional risk, in contrast to a mere failure to reduce or eliminate existing risk. The asymmetry between omission and commission is familiar in legal doctrine, and its impact on judgments of responsibility has been confirmed by psychological research (Ritov and Baron, forthcoming). The asymmetry affects both blame and regret after a mishap, and the anticipation of blame and regret, in turn, could affect behavior.

A moral attitude is involved in another situation where huge discrepancies between buying and selling prices have been observed, the evaluation of environmental amenities in cost benefit analyses. Suppose some corporation offers to buy the Grand Canyon and make it into a water park complete with the world's largest water slide. How do we know whether the benefits of this idea exceed its costs? As usual there are two ways to ask the question, depending on what is the status quo. If there is no theme park in the status quo, then people can be asked the minimum amount of money they would accept to agree to add one (WTA). Alternatively, if the corporation currently owns the right, people could be asked how much they would be willing to pay to buy it back and prevent the theme park from being built (WTP). Several surveys have been conducted where the researchers asked both types of questions for such things as clean air and well-maintained public parks. Most studies find that the WTA responses greatly exceed the WTP answers (Cummings, Brookshire, and Schulze, 1986). The difference in typical responses actually does not tell the entire story. As two close observers of this literature note (Mitchell and Carson, 1989, p. 34): "Studies using WTA questions have consistently received a large number of protest answers, such as 'I refuse to sell' or 'I want an extremely large or infinite amount of compensation for agreeing to this,' and have frequently experienced protest rates [outright refusals to answer the question] of 50 percent or more." These extreme responses reflect the feelings of outrage often seen when communities are faced with the prospect of accepting a new risk such as a nuclear power plant or waste disposal facility (Kunreuther, Easterling, Desvousges, and Slovic, forthcoming). Offers of compensation to proposed communities often do not help as they are typically perceived as bribes.[4]

Judgments of Fairness and Justice

An implication of the endowment effect is that people treat opportunity costs differently than "out-of-pocket" costs. Foregone gains are less painful than perceived losses. This perception is strongly manifested in people's judgments about fair behavior. Kahneman, Knetsch and Thaler (1986) present survey evidence supporting this proposition. Samples of the residents of Toronto and Vancouver were asked a series of questions over the telephone about whether they thought a particular economic action was "fair." In some cases, alternative versions of the same question were presented to different groups of respondents. For each question, respondents were asked to judge whether the action was completely fair, acceptable, somewhat unfair, or very unfair. In reporting the results the first two categories were combined and called "acceptable" and the last two combined and called "unfair." Perceptions of fairness strongly depended on whether the question was framed as a reduction in a gain or an actual loss. For example:

> Question 1a. A shortage has developed for a popular model of automobile, and customers must now wait two months for delivery. A dealer has been selling these cars at list price. Now the dealer prices this model at $200 above list price.

N = 130 Acceptable 29 percent Unfair 71 percent

> Question 1b. A shortage has developed for a popular model of automobile, and customers must now wait two months for delivery. A dealer has been selling these cars at a discount of $200 below list price. Now the dealer sells this model only at list price.

N = 123 Acceptable 58 percent Unfair 42 percent

---

[4] This is a situation in which people loudly say one thing and the theory asserts another. It is of interest that the practitioners of contingent valuation elected to listen to the theory, rather than to the respondents (Cummings, Brookshire and Schulze, 1986). The accepted procedure uses WTP questions to assess value even in a context of compensation, relying on the theoretical argument that WTP and WTA should not be far apart when income effects are small.

Imposing a surcharge (which is likely to be judged a loss) is considered more unfair than eliminating a discount (a reduction of a gain). This distinction explains why firms that charge cash customers one price and credit card customers a higher price always refer to the cash price as a discount rather than to the credit card price as a surcharge (Thaler, 1980).

The different intensity of responses to losses and to foregone gains may help explain why it is easier to cut real wages during inflationary periods:

> Question 2a. A company is making a small profit. It is located in a community experiencing a recession with substantial unemployment but no inflation. The company decides to decrease wages and salaries 7 percent this year.
>
> $N = 125$ Acceptable 37 percent Unfair 63 percent
>
> Question 2b. A company is making a small profit. It is located in a community experiencing a recession with substantial unemployment and inflation of 12 percent. The company decides to increase salaries only 5 percent this year.
>
> $N = 129$ Acceptable 78 percent Unfair 22 percent

In this case a 7 percent cut in real wages is judged reasonably fair when it is framed as a nominal wage increase, but quite unfair when it is posed as a nominal wage cut.

The attitudes of the lay public about fairness, which are represented in their answers to these fairness questions, also pervade the decisions made by judges in many fields of the law. Supreme Court Justice Oliver Wendell Holmes (1897) put the principle this way: "It is in the nature of a man's mind. A thing which you enjoyed and used as your own for a long time, whether property or opinion, takes root in your being and cannot be torn away without your resenting the act and trying to defend yourself, however you came by it. The law can ask no better justification than the deepest instincts of man."

Cohen and Knetsch (1990) showed that this principle, embodied in the old expression that "possession is nine tenths of the law," is reflected in many judicial opinions. For example, in tort law judges make the distinction between "loss by way of expenditure and failure to make gain." In one case, several bales fell from the defendant's truck and hit a utility pole, cutting off power to the plaintiff's plant. The plaintiff was able to recover wages paid to employees which were considered "positive outlays" but could not recover lost profits which were merely "negative losses consisting of a mere deprivation of an opportunity to earn an income" (p. 18). A similar distinction is made in contract law. A party that breaches a contract is more likely to be held to the original terms if the action is taken to make an unforeseen gain than if it is taken to avoid a loss.

Commentary

It is in the nature of economic anomalies that they violate standard theory. The next question is what to do about it. In many cases there is no obvious way to amend the theory to fit the facts, either because too little is known, or because the changes would greatly increase the complexity of the theory and reduce its predictive yield. The anomalies that we have described under the labels of the endowment effect, the status quo bias and loss aversion may be an exceptional case, where the needed amendments in the theory are both obvious and tractable.

The amendments are not trivial: the important notion of a stable preference order must be abandoned in favor of a preference order that depends on the current reference level. A revised version of preference theory would assign a special role to the status quo, giving up some standard assumptions of stability, symmetry and reversibility which the data have shown to be false. But the task is manageable. The generalization of preference theory to indifference curves that are indexed to reference level is straightforward (Tversky and Kahneman, 1991). The factors that determine the reference point in the evaluations of outcomes are reasonably well understood: the role of the status quo, and of entitlements and expectations are sufficiently well established to allow these factors to be used in locating the relevant reference levels for particular analyses.

As Samuelson and Zeckhauser noted, rational models that ignore the status quo tend to predict "greater instability than is observed in the world" (p. 47). We have added the claim that models that ignore loss aversion predict more symmetry and reversibility than are observed in the world, ignoring potentially large differences in the magnitude of responses to gains and to losses. Responses to increases and to decreases in price, for example, might not always be mirror images of each other. The possibility of loss aversion effects suggests, more generally, that treatments of responses to *changes* in economic variables should routinely separate the cases of favorable and unfavorable changes. Introducing such distinctions could improve the precision of predictions at a tolerable price in increased complexity.

After more than a decade of research on this topic we have become convinced that the endowment effect, status quo bias, and the aversion to losses are both robust and important. Then again, we admit that the idea is now part of our endowment, and we are naturally keener to retain it than others might be to acquire it.

■ *The authors wish to acknowledge financial support from Fisheries and Oceans Canada, the Ontario Ministry of the Environment, the Russell Sage Foundation, and Concord Capital Management.*

References

Cohen, David, and Jack L. Knetsch "Judicial Choice and Disparities Between Measures of Economic Values," Simon Fraser University Working Paper, 1990.

Coursey, Don L., John L. Hovis, and William D. Schulze, "The Disparity Between Willingness to Accept and Willingness to Pay Measures of Value," *The Quarterly Journal of Economics*, 1987, *102*, 679–90.

Cummings, Ronald G., David S. Brookshire, and William D. Schulze, eds., *Valuing Environmental Goods*. Totowa, NJ: Rowman and Allanheld, 1986.

Hartman, Raymond, Michael J. Doane, and Chi-Keung Woo, "Consumer Rationality and the Status Quo," *Quarterly Journal of Economics*, forthcoming.

Hershey, John, Eric Johnson, Jacqueline Meszaros, and Matthew Robinson, "What Is the Right to Sue Worth?" Wharton School, University of Pennsylvania, June 1990.

Holmes, Oliver Wendell, "The Path of the Law," *Harvard Law Review*, 1897, *10*, 457–478.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler, "Fairness As a Constraint on Profit Seeking: Entitlements in the Market," *American Economic Review*, 1986, *76*, 728–41.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem," *Journal of Political Economy*, December 1990, *98*, 1325–1348.

Kahneman, Daniel, and Amos Tversky, "Choices, Values and Frames," *American Psychologist*, April 1984, *39*, 341–350.

Knetsch, Jack L., "The Endowment Effect and Evidence of Nonreversible Indifference Curves," *American Economic Review*, 1989, *79*, 1277–1284.

Knetsch, Jack L., "Derived Indifference Curves," working paper, Simon Fraser University, 1990.

Knetsch, Jack L., and J. A. Sinden, "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value," *Quarterly Journal of Economics*, August 1984, *99*, 507–521.

Knetsch, Jack L., and J. A. Sinden, "The Persistence of Evaluation Disparities," *Quarterly Journal of Economics*, 1987, *102*, 691–695.

Knez, Peter, Vernon Smith, and Arlington W. Williams, "Individual Rationality, Market Rationality, and Value Estimation," *American Economic Review*, May 1985, *75*, 397–402.

Kunreuther, Howard, Douglas Easterling, William Desvousges, and Paul Slovic, "Public Attitudes Toward Citing a High Level Nuclear Waste Depository in Nevada," *Risk Analysis*, forthcoming.

Lowenstein, George, and Daniel Kahneman, "Explaining the Endowment Effect," working paper, Department of Social and Decision Sciences, Carnegie Mellon University, 1991.

Mitchell, Robert C., and R. T. Carson, *Using Surveys to Value Public Goods*. Washington, D.C.: Resources for the Future, 1989.

Ritov, Rita, and Jonathan Baron, "Status-quo and Omission Biases," *Journal of Risk and Uncertainty*, forthcoming.

Samuelson, William, and Richard Zeckhauser, "Status Quo Bias in Decision Making," *Journal of Risk and Uncertainty*, 1988, *1*, 7–59.

Thaler, Richard, "Toward a Positive Theory of Consumer Choice," *Journal of Economic Behavior and Organization*, 1980, *1*, 39–60.

Tversky, Amos, and Daniel Kahneman, "Loss Aversion and Riskless Choice: A Reference Dependent Model," *Quarterly Journal of Economics*, 1991.

Viscusi, W. Kip, Wesley A. Magat, and Joel Huber, "An Investigation of the Rationality of Consumer Valuations of Multiple Health Risks," *RAND Journal of Economics*, 1987, *18*, 465–79.

ATTACHMENT D

Journal of Economic Behavior and Organization 17 (1992) 131–139. North-Holland

Preferences and nonreversibility of indifference curves*

Jack L. Knetsch

School of Resource and Environmental Management, Simon Fraser University, Burnaby, B.C. V5A 1S6, Canada

Received July 1990, final version received March 1991

New evidence indicating that people commonly demand more to give up an entitlement than they would pay to acquire the same good suggests that reversible indifference curves may not accurately reflect consumer preference orderings. This paper presents the results of two direct experimental tests of such reversibilities. The tests used two commodity sets, two different experimental designs, and both student and non-student participants. The results of both experiments showed large valuation disparities, consistent with other tests of an endowment effect, and nonreversibility of indifference curves.

1. Introduction

The assumption that indifference curves are reversible, that the rate at which people will substitute one good for another is independent of their initial entitlements and the direction of exchange offers, underlies much of contemporary demand theory and is at least implicit in a vast array of applied economics. However, the validity of this commonly postulated property of people's preferences is brought into question by the growing evidence that people commonly demand much more to give up an entitlement than they are willing to pay to acquire a fully commensurate one [Kahneman and Tversky (1979), Gordon and Knetsch (1979), Thaler (1980), Knetsch and Sinden (1984), Brookshire and Coursey (1987), and Kahneman, Knetsch and Thaler (1990)].

Conventional principles of economic theory hold that an indifference curve shows the rate at 'which one good can be substituted for another in consumption without changing the level of satisfaction or well-being in consumption' [Freeman (1983, p. 75)], and 'The slope of an indifference curve measures ... the number of units of one of the goods necessary to

---

*This research was supported, in part, by Fisheries and Oceans Canada, Environment Canada, and the Ontario Ministry of the Environment. It has benefited from suggestions and comments of Daniel Kahneman, Peter Kennedy, Andrew Mueller, Paul Slovic and Richard Thaler, and the suggestions and research assistance of Sandra Broughton and Sheila Fagnan.

0167-2681/91/$03.50 © 1991—Elsevier Science Publishers B.V. All rights reserved

replace one unit of the other good in order to maintain the same total satisfaction' [Clower, Graves and Sexton (1988, p. 93)]. However, to the extent that people value losses more than foregone gains, as the evidence from a wide variety of controlled tests suggests, they would demand greater compensation to give up a good than they would be willing to pay to obtain an equal amount of the same good. The terms of acceptable exchanges will then vary depending on the direction of proposed trades indicating that indifference curves would not be reversible.

The purpose here is to report the results of two more direct tests of the reversibility of indifference curves. The first uses an experimental method, initially demonstrated by MacCrimmon and Toda (1969), to empirically derive individuals' indifference curves between a single good, pens, and money. The second uses a simpler incentive compatible experimental design to test reversibility between two consumption goods.

2. Derived indifference curves experiment

In the first experiment, indifference curves were experimentally derived for each participant following the procedures used by MacCrimmon and Toda (1969).[1] In their experiments, all subjects first received a sum of money, and then each was offered a series of trades of specific numbers of ball point pens (or pastries, in a second experiment) for part of the money. Participants specified which of the offers of pens they preferred to keeping all of the money and which they did not prefer. After a number of options were presented and the preferences recorded, an indifference curve was drawn over the boundary between combinations of pens and money which each individual had indicated they preferred to their original entitlement and those which they rejected in favour of their initial holding of money. In order to determine a series of indifference curves for each individual, different amounts of money were given to each as initial endowments.

The method used to obtain the results reported here followed the same procedure. Each participant was first given an initial entitlement and was then offered a series of successive choices of keeping this endowment or giving up part of it in exchange for another good. Participants' indicated preferences were recorded after each choice.

The important extension in this experiment was that half of the individuals were initially given pens and then offered varied amounts of money in exchange for some or all of the pens, and the other half were first given cash and then offered different numbers of pens in exchange for varying amounts of the money. Also, as this was a test for any effect of the initial endowment and the direction of subsequent exchange offers on the slope of the resulting indifference curve, only one curve was estimated for each individual.

---

[1] Although the method has apparently not often been used, MacCrimmon and Toda's results continue to be widely cited, especially in micro-economics textbooks [see, for example, Hirshleifer (1988, p. 70), Truett and Truett (1984, p. 63), and Mansfield (1988, pp. 68–69)], likely because of their empirical verification of assumed characteristics of consumer behaviour – 'the extent to which the curves satisfy the commonly postulated properties of indifference curves' [MacCrimmon and Toda (1969, p. 433)].

The sample of participants consisted of 20 women, all of whom were interviewed at a public swimming pool while they waited for their children to finish their activities at the pool. Each of the participants was given either $4.50 or five reasonably high quality ball point pens for agreeing to complete a short questionnaire. Participants were told that the form of their payment had been determined randomly, and that about half of them received money and the other half pens.

After each participant received the randomly determined payment of either pens or money and the questionnaire was completed, each individual was given either a series of offers of from one to ten pens in exchange for giving up some of the money or a series of offers of from one to ten dollars in exchange for pens, depending on her initial entitlement. As in the MacCrimmon and Toda experiments, each participant was also told that she would actually receive one of the combinations of pens and money which she had indicated was preferred to the initial payment. The pay-off was selected randomly from the set of all combinations that each participant stated she preferred to her original endowment.

Half of the participants in each sample group were initially offered an exchange that would take nearly all of their original entitlement of money or pens. The first offer made to the other half of each group was a trade that would take only a small portion of the payment. This was done to counter any possibility of an anchoring effect that might bias reactions to subsequent offers in accord with the response to the first offer. No such effect was in fact observed.

The procedure is illustrated in fig. 1, where the acceptances and rejections of offers presented to one of the participants are plotted. This person had received five pens as payment for completing the questionnaire. The first offer asked if the participant would prefer to keep the five pens or to give up all of them for $7.00. She chose to accept the $7.00, and this is indicated by an 'x' in the figure. A second offer of $4.00 for three of the pens was refused, and this is indicated by an 'o' in the figure. Twelve further offers of various amounts of money for some of the pens were then made, with each put in the form of a question asking if the new combination was or was not preferred to keeping the original five pens. The accepted and rejected offers are indicated by the 'x' and 'o' plots. The curve of presumed indifference between money and pens was then drawn between the combinations of accepted and rejected offers.

An average of 17.8 separate offers were made to the individual participants.

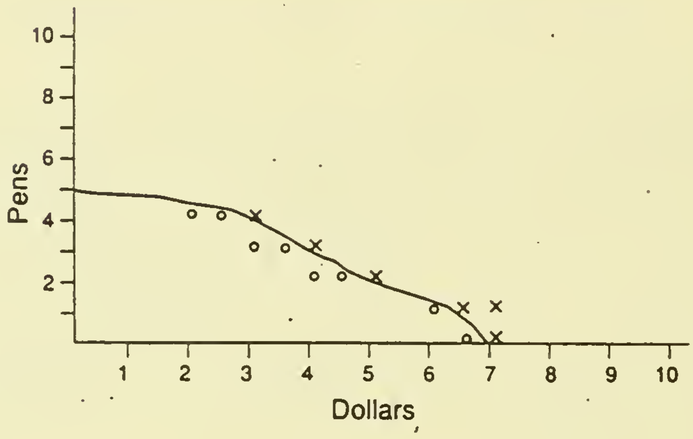


Fig. 1. Accepted and rejected trade offers of one individual.

The pay-off scheme used by MacCrimmon and Toda, and in the present experiment, introduces a potential incentive for strategic behaviour that could exaggerate the apparent reluctance to give up the initial entitlements. Having been informed that they would actually receive one of their preferred combinations of money and pens, respondents might increase the desirability of the pool of possible payoff combinations by comparing each subsequent offer to the most preferred of the ones that they had previously accepted rather than to their original entitlement. A strategy of refusing an offer that is preferred to the original endowment, but is turned down because it is less desirable than a previously accepted trade, could produce a bias in the estimated curves.

The actual responses indicate, however, that such strategic behaviour was likely neither a consistent nor an important motivation. The order of only the first three offers made to each respondent was recorded, but even in these initial choices 35 percent of the subjects had already acted in direct contradiction to the strategy by accepting offers that, given positive evaluations of the two goods, could not have been superior to offers accepted earlier. Slightly more than half of the remaining participants rejected offers for pens or dollars that far exceeded the number that respondents in the opposite group indicated they would not give up for more money or pens. In terms of fig. 1, these individuals turned down offers of nine or ten pens in exchange for three or four dollars even through all people initially given five pens preferred keeping them unless compensated with far larger sums. In the remaining cases – 30 percent – there was no evidence either consistent with or contrary to choices being influenced by strategic behaviour incentives.

The indifference curves derived for individuals in the two groups differed dramatically. Two of the ten people in each group indicated that they would

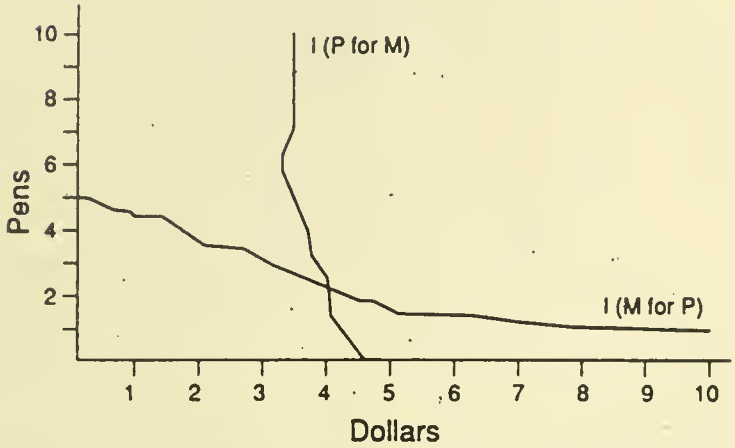


Fig. 2. Average indifference curves resulting from offers of money for pens (M for P) and offers of pens for money (P for M).

not be willing to give up any of their original entitlement for the money or pens offered in exchange. All of the remaining participants in both groups were willing to give up some of their original endowment, but a reluctance to exchange was strong and consistent.

A summary of the results, in the form of the average indifference curve for all members of the two groups, is given in fig. 2. These average curves are reasonbably characteristic of the individual curves for each group. The curve labelled I(P for M) gives the results from trade offers of pens for money. This curve, and the ten individual curves it summarizes, are very similar to those obtained by MacCrimmon and Toda when they offered pens for money to the seven participants in their experiment (1969, pp. 442–443).

The other curve in fig. 2, labeled I(M (for P), is the average indifference curve for the ten individuals, randomly selected from the group of participants, who were offered varying amounts of money to give up pens. The slope of this curve is very different from that of the other indifference, curve.

While there is no reason to expect each individual to have identical preferences, the results of this test indicate large systematic differences in the choices. The combinations of goods that are considered to be of equal worth, and between which people are willing to make trades, differed widely depending on the direction of the trade offers. In this test, the indifference curves are clearly not reversible.

3. Exchange of two goods experiment

The second experimental test of the reversibility of indifference curves involved simple exchange offerings of giving up one entitlement to acquire an alternative – a somewhat more stringent variation of a design used earlier [Knetsch (1989)] which eliminates the strategic behaviour possibility in the first experiment. One half of the participants were offered a pen in exchange for a mug and the other half were offered a mug in exchange for a pen.

The experiment was conducted over a three day period in six separate tutorial groups, of from 11 to 22 people each, of a large pre-principles economics class at Simon Fraser University. All participants were asked to take part in a research project that involved answering a short questionnaire. In three of the groups each participant was then given a survey form and a coffee mug with sketch of the University on it. Participants in the other three groups were given a questionnaire and a ball point pen in a box with a clear side permitting it to be seen.[2] All were told that the good was theirs to keep.

After the questionnaires were completed and collected participants were again informed that the good was theirs to keep, but they were then also told that they could give it up in exchange for the other good – either a pen or a mug depending on which was given initially – plus five cents. Quantities of the alternative good, and a pile of five cent pieces, were then placed in full view of the participants. They were asked to indicate if they wished to make the exchange. All exchanges were carried out immediately, with no uncertainties or delays.

Reversibility of indifference curves requires that valuations of the two goods by an individual are independent of the initial entitlements – the independence axiom – and that the relative values are the same for trades in either direction. In terms of the simple choices presented in this experiment, participants' evaluation of a mug relative to a pen should be the same whether the choice is in the form of giving up a mug for a pen or a pen for a mug – 'since a receipt foregone of a given amount is the equivalent of a payment of the same amount' [Coase (1960, p. 7)]. Consequently, if an individual's valuations of losses and opportunity costs are equivalent, as conventional views of reversible indifference curves assume, roughly the same proportion of people should prefer a mug to a pen, or a pen to a mug, regardless of their initial entitlement.

If anything, dominance would lead to more participants in all groups giving up their initial entitlements as five cents is acquired in each trade – a good plus five cents dominates an identical good without the five cents and this should show in the resulting choices. However, the results show a strong endowment effect of losses being systematically valued more than gains even though the gains included five cents that was not included in any of the entitlements given up. The preferences indicated by participants' choices between keeping one good or trading a mug or a pen for the other, plus five cents, were:

---

[2] The coffee mugs were available in the University bookstore. Individuals receiving pens were instructed to leave them in the boxes so that those in possession of the pens had no further information about them than those who were contemplating trading for them.

Start with	Exchange offer	Prefer Pens	Prefer Mugs	N
Mug	Pen + 5c	12%	88%	51
Pen	Mug + 5c	90%	10%	49

The preference for a pen or a mug varied systematically, and significantly (chi-squared is 60.86), depending on the direction of the exchange.[3] A pen was considered to be worth more than a mug by 90 percent of the participants when it was valued as a loss, but a pen was valued more than a mug by only 12 percent of the participants when it was evaluated as a gain.

Pens were preferred to mugs by a margin of about 9 to 1; but mugs were also preferred to pens by about the same ratio. The preference for one good over the other was very much a function of which was gained and which was given up, and was not a result of a consistent preference for one good over the other, notwithstanding the symmetry of the exchange offers and the five cents added to the offers.

4. Discussion

Both experiments provided a test of the equivalence in people's valuation of gains and losses – that they value A over B or B over A regardless of which good is offered in exchange for the other. Such symmetry is required for the traditional assertion of reversibility of indifference curves.

The designs of these two tests varied greatly, although both offered one half of the participants exchanges in one direction and the other half trades in the opposite direction. In both cases there were no income, or wealth, effects that could influence the valuation of gains and losses, nor any income constraints that could hinder an individual's ability to make a desired exchange. Although a strategic bias is possible in the MacCrimmon–Toda design experiment, this is eliminated in the second experiment. One experiment used student subjects, the other was conducted with non-student participants.

---

[3] All participants were asked not to discuss the exercise with others in later tutorial groups who subsequently might take part in the experiment. To the extent that people in later sections were apprised of the experiment, it would presumably have alerted them to the exchange possibility and thereby weakened their focus on the initial entitlement as their reference position. In that event the endowment effect would have been reduced.

The results of the two tests were very similar: both indicate that losses matter more to people than gains. If the participants followed the usual textbook assumption 'that the choices they make are preferred to the choices that they could have made' [Varian (1990, p. 120)], then the valuations of goods depended to a large degree on the reference position at the time of the exchange offer and the direction of change from this point. The first experiment demonstrated that the individuals were far from indifferent between the two initial entitlements as participants in one group turned down offers far in excess of those kept by people in the other group. And while people in the second experiment might have been about equally pleased to have either a pen or a mug, other parallel experiments in which similar subjects made bids and offers for similar goods revealed that both buyers and sellers display a wide range of values, indicating that in the absence of an endowment effect there would be enough rents to produce significant gains from trade [Kahneman, Knetsch and Thaler (1990)].

These results are in full accord with other tests of the endowment effect and nonreversibility of indifference curves that used other experimental and survey designs [Knetsch (1989)]. They are also consistent with other findings of large valuation disparities that are independent of strategic behaviour, income effects, transaction costs, or repetition [Kahneman, Knetsch and Thaler (1990)]. The valuation differences and nonreversibility of indifference curves appear on present evidence to be far more pervasive characteristics of people's preference orderings than the normally assumed equivalence in valuations of gains and losses and reversibility of curves.

To the extent that indifference curves are not reversible, predictions and estimates based on traditional assumptions of equivalence and reversibilities may be seriously in error. For example, gains from trade will be overstated, too many voluntary exchanges will be predicted, and welfare losses will be under estimated. The disparities will also lead to a reluctance to yield concessions and to judgments that actions that benefit one party at the expense of another are unfair [Kahneman, Knetsch and Thaler (1986)].

References

Brookshire, David S. and Don L. Coursey, 1987, Measuring the value of a public good: An empirical comparison of elicitation procedures, American Economic Review 77, no. 4, 554–566.

Clower, Robert W., Philip E. Graves and Robert L. Sexton, 1988, Intermediate microeconomics (Harcourt Brace Jovanovich, Orlando, FL).

Coase, Ronald H., 1960, The problem of social cost, The Journal of Law and Economics 3, 1–44.

Freeman, A. Myrick, 1983, Intermediate microeconomic analysis (Harper & Row, New York).

Gordon, Irene M. and Jack L. Knetsch, 1979, Consumer's surplus measures and the evaluation of resources, Land Economics 55, no. 1, 1–10.

Hirshleifer, Jack, 1988, Price theory and applications, 4th ed. (Prentice-Hall, Englewood Cliffs, NJ).

Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler, 1986, Fairness as a constraint on profit seeking: Entitlements in the market, The American Economic Review 76, no. 4, 728–741.

Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler, 1990, Experimental tests of the endowment effect and the Coase theorem, Journal of Political Economy 90, no. 6, 1325–1348.

Kahneman, Daniel and Amos Tversky, 1979, Prospect theory: An analysis of decisions under risk, Econometrica 47, no. 2, 263–291.

Knetsch, Jack L., 1989, The endowment effect and evidence of non-reversible indifference curves, The American Economic Review 79, no. 5, 1277–1288.

Knetsch, Jack L. and John A. Sinden, 1984, Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value, The Quarterly Journal of Economics XCIX, no. 3, 507–521.

MacCrimmon, Kenneth R. and M. Toda, 1969, The experimental determination of indifference curves, Review of Economic Studies 36, 433–451.

Mansfield, Edwin., 1988, Micro-economics, theory and applications, 6th ed. (Norton, New York).

Thaler, Richard H., 1980, Toward a positive theory of consumer choice, Journal of Economic Behavior and Organization 1, no. 1, 39–60.

Truett, Lila J. and Dale B. Truett, 1984, Intermediate microeconomics (West Publishing, Saint Paul).

Varian, Hal R., 1990, Intermediate microeconomics, A modern approach, 2nd ed. (W.W. Norton & Co., New York).

ATTACHMENT E

ASYMMETRIC VALUATION OF GAINS AND LOSSES AND PREFERENCE ORDER ASSUMPTIONS OF ECONOMICS[1]

Jack L. Knetsch

Professor of Economics and Resource Management
School of Resource and Environmental Management
Simon Fraser University
Burnaby, British Columbia
Canada V5A 1S6

Abstract

Recent findings of large differences between people's valuations of gains and commensurate losses imply that common preference order axioms of completeness, transitivity, and dominance, on which economic analyses and predictions of consumer behaviour are largely based, may not be consistent with the actual choices of many consumers.

The results of an explicit experimental test of the preference axioms, using real exchanges and consequent weighing of actual gains and losses, are reported. The findings show consistent differences in valuations of goods and money depending on whether the entitlements are being acquired or given up, and provide direct evidence of preference assumption violations.

---

[1] This research was supported, in part, by the Social Science and Humanities Research Council of Canada, The Science Council of British Columbia, and the Ontario Ministry of the Environment, and has benefited from the suggestions of Daniel Kahneman and Richard Thaler and the comments and assistance of Sheila Fagnan.

ASYMMETRIC VALUATION OF GAINS AND LOSSES AND PREFERENCE ORDER ASSUMPTIONS OF ECONOMICS

Completeness, dominance, transitivity, independence and related axioms of "consistent" preference orderings provide major supports for much of the contemporary application of demand theory and practice of economics[2] -- "economists today simply say that individuals tend to make consistent choices, that the[se] Laws of Preference ..., while undoubtedly idealized, are a good approximation of actual behavior" (Hirshleifer, 1988, p. 59).

However, recent findings that people commonly demand more to give up an entitlement than they are willing to pay to acquire or keep an identical one suggests that these common preference assumptions, which "are so fundamental that we can refer to them as 'axioms' of consumer theory" (Varian, 1990, p. 35), may not usefully describe actual preferences. The evidence from a wide array of recent survey studies, controlled experimental tests, and observations of common economic behaviour, reveals large and pervasive asymmetric valuations of gains and losses and indicates that individuals commonly prefer to keep a good A rather than give it up for another good B but also prefer to keep B rather than exchange it for A if given the opposite choice. Such concurrent preference orderings are inconsistent with dominance, independence, completeness, and transitivity criteria -- the case, for example, if people prefer A over B when asked to give up A for B, prefer not to give up B for C, and yet prefer not to give up C to acquire A.

---

[2] It is commonly assumed that "all conceivable baskets of commodities can be consistently ranked in order of preference by the individual" (Hirshleifer, 1988, p. 49); "if one option is better than another in one state and at least as good in all other states, the dominant option should be chosen" (Tversky and Kahneman, 1986, p. 69); "if he [or she] prefers A to B and prefers B to C, then he [or she] always prefers A to C" (Frank, 1991, p. 69); and "choices that people plan to make in one state of nature should be independent from the choices they plan to make in other states of nature" (Varian, 1990, p. 216).

The results of past tests of preference orderings imply that such violations may well be common, but these examinations were based primarily on hypothetical choices among risky prospects (Tversky, 1969; Tversky and Kahneman, 1986; Thaler, 1987; Loomes, 1991; and Loomes and Taylor, 1992). The purpose here is to report the results of an explicit test of common preference order assumptions which used a series of riskless real exchange offers to individuals to give up one of three entitlements to acquire one of the alternatives. The test design eliminated indifference as a possible explanation for the results and provided comparisons both between goods and between goods and money.

The Experiment

The test was conducted over a three day period in 18 separate tutorial groups, of from 11 to 22 people each, of the same large pre-principles economics class at Simon Fraser University. All participants were asked to take part in a research project that involved answering a short (and unrelated) questionnaire. All of the individuals in each group were then given a survey form and one of three goods: a coffee mug with a sketch of the University on it, a ball point pen, or $2.[3] All participants were told that the good given to them -- either a mug, a pen, or $2 -- was theirs to keep.

After the questionnaires were completed and collected, participants were again informed that they could keep the good they had been given, but they were then also told that they could give it up in exchange for a second good plus five cents. Quantities of the alternative good -- enough to complete an exchange for all of the participants in each group -- plus a pile of five cent pieces, were then placed in full view of the participants and they were asked to indicate if they wished to make a trade. All exchanges were carried out immediately.

All six of the possible combinations of initial entitlement and exchange offers of the three goods -- a mug+5c for a pen, a mug+5c for $2, a pen+5c for a mug, a pen+5c for $2, $2+5c for a mug, and $2+5c for a pen -- were carried out over the 18 tutorial groups, with each initial-entitlement/subsequent-offer combination presented to about 50 individuals in three tutorials.[4]

---

3. The coffee mugs were available in the University bookstore. Individuals receiving pens were told to leave them in their display boxes so that those in possession of the pens had no further information about their quality or performance than those contemplating trading for them.

4. All participants were asked not to discuss the exercise with others who might take part in later tutorial groups. To the extent that people in later sections were apprised of the experiment it would presumably have alerted them to the exchange possibility and thereby weakened their focus on the in-

Results

Conventional economic assumptions provide a strong prediction of the expected choices. These traditional axioms suggest that the evaluation of one good relative to another should be the same whether the choice is in the form of giving up one for the second or giving up the second for the first -- as in the common assertion that "a receipt foregone is the equivalent of a payment of the same amount" (Coase, 1960, p. 7). Therefore, at a minimum the proportions of people preferring one or the other of each pair of goods should be roughly the same, subject only to experimental variation, regardless of the direction of the possible trade. And to the extent that the five cents added to each offered good meant that these on average dominated the original holdings -- a pen plus five cents is better than a pen alone -- the numbers of individuals making trades should then exceed the numbers choosing to keep their initial good.

The results strongly contradicted the outcomes expected on the basis of the usual preference assumptions. Comparisons among all three of the pair-wise valuations of the three goods revealed an overwhelming preference for keeping the initial good, regardless of what it was, over giving it up for either of the two alternative entitlements offered in exchange. In each case

---

itial entitlement as their reference position and consequently reduced the endowment effect.

participants indicated that they valued the losses of their initial entitlement more than the foregone gains of not acquiring an alternative.

The preference orderings revealed by the choices made by the 100 individuals choosing between mugs and pens were:

Start with	Exchange Offer	Prefer Pens	Prefer Mugs
Mug (N=51)	Pen + 5c	12%	88%
Pen (N=49)	Mug + 5c	90%	10%

The ranking of preferences between a pen and a mug varied greatly, and significantly (chi-squared 60.86, p<.001), depending on the direction of the exchange offers. Eighty-eight per cent of the participants revealed a preference for a mug over a pen when the mug was valued as a loss and the pen as a gain, but a mug was preferred to a pen by only 10 per cent of the participates when the mug was evaluated as a gain and the pen as a loss.

The preference variations were much the same for valuations of a mug relative to $2:

Start with	Exchange Offer	Prefer Mugs	Prefer $2
Mug (N=58)	$2 + 5c	97%	3%
$2 (N=44)	Mug + 5c	18%	82%

The difference in the valuations is again large and significant (chi-squared 65.74).

The pattern of preferences did not change for the remaining pair of entitlements, money and pens:

Start with	Exchange Offer	Prefer Pens	$2
Pen (N=51)	$2 + 5c	82%	18%
$2 (N=49)	Pen + 5c	14%	86%

The magnitude of the gain/loss disparity, and the significance (chi-squared 46.33) was about the same as in the other two comparisons.

Given the random distribution of the initial entitlements across all of the 302 participants, about half of them would be expected to make an exchange if preferences were in reality independent of entitlements as implied by the traditional preference axioms. However, rather than the expected 151 trades over all of the exchange possibilities, only 37 participants, or 12.3 per cent, agreed to give up their initial entitlement for an alternative. This failure to give up initial entitlements occurred in spite of the symmetry of the initial and the offered entitlements, and all participants being on average slightly better off by trading and receiving an added 5 cents for doing so.

A similar lack of voluntary trades could also have been observed if nearly all of the participants in each group had been completely indifferent between having one or the other of the pair of entitlements offered to them. In that case, they could not make themselves better off by making an exchange and this would lead them to stay with their initial entitlement. This would have needed to be the case not only for exchange possibilities involving the two goods used in the experiment, pens and mugs, but also for comparisons between each of these goods and money.

The five cents included in all of the offers essentially precluded indifference giving rise to the observed reluctance to trade. This inclusion assured that the entitlements offered in exchange dominated the entitlements given up because they were identical in all respects except that the exchange offers included five cents -- a mug plus five cents, for example, is marginally better and certainly not worse than a mug alone. An individual who might be indifferent between a pen and a mug plus five cents, presumably must prefer a pen plus five cents to a mug. To the extent that people are presumed to favour the dominant entitlement in each pair-wise comparison, more, and not fewer, of them should choose the good offered in exchange.

Further, the evidence of wide ranges of valuations and systematic large differences between gains and losses, from roughly parallel experiments using similar goods and comparable individuals, suggests that indifference is a very unlikely explanation for the present results. For example, the average sum demanded to give up a lottery ticket was found to be about four times larger than the average amount that individuals would pay to acquire an identical entitlement (Knetsch and Sinden, 1984). In another incentive compatible experiment in which participants were given either money or chocolate bars and made bids and offers to give up or acquire chocolate bars, the people starting with the chocolate valued the bars at a little over twice as much as the people who were given money (Knetsch, 1989). And in a series of evaluation exercises involving repeated bids and offers for goods such as pens, mugs, and cardboard binoculars, the ratios of the median final valuations of losses and gains ranged from about two to nearly seven (Kahneman, Knetsch and Thaler, 1990).

The results are also not attributable to income, or wealth, effects, or to income constraints. The experimental design, which endowed each participate with an entitlement and gave the person the opportunity to exchange it for another one, eliminated these possible sources of valuation difference.

Although transaction costs -- including the mental effort of comparing values -- can give rise to inhibitions to trade, the results of this and other studies suggest that such costs do not explain this or many other observed cases of apparent valuation differences. For example, in a series of experimental tests, people were first allowed to trade entitlements with values prescribed for them, and subsequently were then given the opportunity to trade real goods that they valued themselves (Kahneman, Knetsch and Thaler, 1990). Even though the trading rules, and presumably any transaction costs, were identical in the two markets, participants traded the induced value entitlements fully in accord with standard assumptions, but showed widely divergent valuations of gains and losses and a strong and consistent reluctance to trade the real goods.

The inference of valuation disparities is based on between subject comparisons, in this experiment, as no individual was presented with more than one choice and therefore each expressed only a single preference between giving keeping a good A or giving it up for another good B. However, similar valuation disparities have been found in within subject design studies, when the same individuals valued both gains and losses in making hypothetical choices (e.g. Hammack and Brown, 1974) or in buying and selling entitlements in real exchanges (Kachelmeier and Shehata, 1992). Further, the results of this between subject design experiment can be used to infer preference orderings of a representative individual across all of the choice possibilities. That is, on the basis of the results of this experiment, a representative individual faced with a choice between a pen and a mug is far more likely to prefer a pen when asked to choose between giving up a pen to acquire a mug than when asked to acquire a pen by giving up a mug.

The disparity between valuations of gains and losses appears to be a common characteristic of people's preferences. And it is this pervasive difference, rather than transaction costs or wealth effects, that seems to be the major reason for the observed reluctance of people to give up original, or reference, entitlements to acquire an alternative. Parties simply perceive fewer gains from trade than are generally presumed to be available on the basis of assumed valuation equivalence.

Discussion

The actual choices of the participants in this real exchange experiment clearly did not conform to those expected on the basis of normal preference axioms. Further, the patterns were the same for choices between two consumption goods, and those between money and either of the goods.

Contrary to the completeness axiom, these individuals indicated neither consistent indifference nor a consistent preference for one or the other good in any of the pairings. The outcomes of the comparisons were not independent of the basis for the comparison. Nor did the value of one good dominate the value of another. In one reference position pens were preferred to mugs (plus five cents) by a margin of about 9 to 1; in the alternative reference position mugs were preferred to pens (plus five cents) by about the same ratio. The results were about the same for the other two pairings.

The observed preference orderings are also inconsistent with transitivity assumptions. People greatly preferred pens over mugs plus 5c, mugs over $2.05, but $2 over pens plus 5c. And similarly they indicated valuations to support all other possible preference orderings.

Contrary to commonly assumed behaviour, the preference orderings were not at all consistent. One good was not consistently preferred to either of the others, again not withstanding the symmetry of the choice options and the five cents added to the entitlement offered in exchange for the initial reference entitlement. In aggregate, the participants remained at what could be read as nominally inferior positions, but were in reality superior ones.

The violations of transitivity and other axioms indicated by the results of this experiment occur because gains or losses of a good have been implicitly assumed in the axioms to be viewed by people as simply having more or less of a good, without the possibility for any disparity in their valuations due to the direction of change. A more accurate accounting of welfare changes would result from an explicit allowance for this difference -- defining changes in terms of their sign (positive or negative) as well as their quantity (Kahneman and Tversky, 1979; and Tversky and Kahneman, 1991). Thus, in the case of the present exercise: a negative one mug is valued more than a positive one pen, a negative pen is worth more than a positive mug, and so forth for the other pairs of choices. Viewed in these terms, the experimental results do not demonstrate violations of transitivity, dominance or any other common axiom. However, such a view of the impact of changes on welfare does require allowing for value to be dependent on the direction of changes from a reference entitlement rather than being independent of entitlement and equivalent in the domains of either gains or losses, as is usually assumed.

Recognition of the disparity in valuation of gains and losses as a pervasive characteristic of preferences appears, first, to lead to more accurate descriptions of behaviour -- a more parsimonious and satisfactory explanation of many "anomalies" than their being attributable to transaction costs of some undefined kind. Second, this characterization offers more useful guidance for policy proposals. For example, reducing transaction costs may frequently be worthwhile, but the impacts of such efforts are likely to be more modest than otherwise expected because of the valuation difference; and policy options or negotiating proposals that result in some people foregoing gains may be more acceptable than others that attempt to reach similar goals by imposing losses on people (Kahneman and Tversky, to appear; and Borges and

Knetsch, 1993). Further, establishing a market for transferable pollution, fishing, or water licenses may increase their reallocation towards more efficient uses, but it probably will not eliminate the consequences of their initial distribution. Similarly, and also contrary to common prescriptions, even partial mitigation of harms may well be preferred to a cash compensation remedy (Knetsch, 1990). And people's judgments of the fairness of price and employment changes, and their acceptance of remedial measures, may be heavily influenced by their perceptions of how losses are imposed as a result of the action (Kahneman, Knetsch and Thaler, 1986).

Losses appear to matter much more than gains to most people, and valuations of goods depends, to a large degree, on the reference position and the direction of change from this point. Without an accounting for this disparity, consumer choices will often violate the usual axioms of preference orderings, and predictions and prescriptions based on them may well be much less useful than ones that allow for this common characteristic of actual behaviour.

References

Borges, Bernhard F.J., and Jack L. Knetsch, (1993), "Valuation of Gains and Losses, Fairness, and Negotiation Outcomes", Simon Fraser University Working Paper.

Coase, Ronald H., (1960), "The Problem of Social Cost", The Journal of Law and Economics, 3, 1-44.

Frank, Robert H., (1991), Microeconomics and Behavior, (McGraw-Hill, New York).

Hammack, Judd, and Gardner M. Brown, (1974), Waterfowl and Wetlands: Toward Bio-Economic Analysis, (Johns Hopkins, Baltimore).

Hirshleifer, Jack, (1988), Price Theory and Applications, 4th ed., (Englewood Cliffs, N.J.: Prentice Hall).

Kachelmeier, Steven J., and Mohamed Shehata, (1992), "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China", The American Economic Review, 82, 1120-41.

Kahneman, Daniel, and Amos Tversky, (1979), "Prospect Theory: An Analysis of Decision Under Risk", Econometrica, 47, 263-291.

Kahneman, Daniel, and Amos Tversky, (to appear), "Conflict Resolution: A Cognitive Perspective", in K. Arrow, et.al., eds., Barriers to Conflict Resolution, (Norton, New York).

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler, (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem", The Journal of Political Economy, 98, 1325-1348.

Kahneman, Daniel, Jack L. Knetsch, and Richard Thaler, (1986), "Fairness as a Constraint on Profit-Seeking: Entitlements in the Market", The American Economic Review, 76, 728-741.

Knetsch, Jack L., (1990), "Environmental Policy Implications of Disparities Between Willingness to Pay and Compensation Demanded Measures of Values", Journal of Environmental Economics and Management, 18, 227-237.

Knetsch, Jack L., (1989), "The Endowment Effect and Evidence of Nonreversible Indifference Curves", The American Economic Review, 79, 1277-1284.

Knetsch, Jack L., and John A. Sinden, (1984), "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value", The Quarterly Journal of Economics, 99, 507-521.

Loomes, Graham, (1991), "Evidence of a New Violation of the Independence Axiom", Journal of Risk and Uncertainty, 4, 91-108.

Loomes, Graham, and Caron Taylor, (1992), "Non-Transitive Preferences Over Gains and Losses", The Economic Journal, 102, 357-65.

Thaler, Richard H., (1987), "The Psychology of Choice and the Assumptions of Economics", in Alvin E. Roth, ed., Laboratory Experimentation in Economics, (Cambridge: Cambridge University Press).

Tversky, Amos, (1969), "Intransitivity of Preferences", Psychological Review, 76, 31-48.

Tversky, Amos, and Daniel Kahneman, (1986), "Rational Choice and the Framing of Decisions", The Journal of Business, 59, S251-S278.

Tversky, Amos, and Daniel Kahneman, (1991), "Loss Aversion in Riskless Choice: A Reference-Dependent Model", The Quarterly Journal of Economics, 1039-1061.

Varian, Hal R., (1990), Intermediate Microeconomics: A Modern Approach, 2nd ed., (New York: W.W. Norton & Co.).

ATTACHMENT F

JUDICIAL CHOICE AND DISPARITIES BETWEEN MEASURES OF ECONOMIC VALUES©

By David Cohen[*] and Jack L Knetsch[**]

An important idea which characterizes law in society is a reluctance to move from the *status quo*. In general, one can argue that legal institutions and legal doctrine are not engaged in the redistribution of wealth from one party to another. This paper explores a possible explanation for that principle. The authors' research suggests that, across a wide range of entitlements and in a variety of contexts, individuals value losses more than foregone gains. The paper argues, as a matter of efficiency, that law and social policy might have developed in a manner consistent with this valuation disparity. Furthermore, this valuation disparity can be transformed into conceptions of fairness, and that as a matter of fariness, legal decisions might have developed in a manner consistent with this fairness norm. In the first part of the paper, the economic and psychological research on the valutation disparity is described in detail. The paper then examines a series of legal doctrines, all of which can be explained by the valuation disparity phenomenon revealed in the experimental data. Cohen and Knetsch conclude that the behaviour of legal institutions and actors can be explained by the valuation disparity.

I. INTRODUCTION

The idea that the legal system should not move wealth from one person to another pervades common law doctrine and reasoning — as Oliver Wendell Holmes stated, "The general principle of our law is that

---

© Copyright, 1992, D. Cohen and J.L Knetsch.

[*] Professor of Law, University of British Columbia.

[**] Professor ~~at the School of~~ of Economics and Resource ~~and Environmental~~ Management, Simon Fraser University. The authors acknowledge the support of the Social Sciences and Humanities Research Council of Canada, the Ontario Ministry of the Environment and the British Columbia Science Council. The authors also wish to thank William Fischel and Stephen Waddams for their helpful comments and further examples.

loss from accident must lie where it falls."[1] Common explanations of that position focus on the political power and class bias of those who make legal decisions and create legal rules.

We propose an alternative interpretation of the historical truism that "losses should lie where they fall." People value actual losses far more than foregone gains, and thus, it is a matter of efficiency,[2] as well as of fairness, to adopt presumptive legal rules which do not direct non-consensual transfers of wealth. We explore the extent to which such disparities in valuations in gains and losses have been incorporated in legal doctrines and principles.

The following two puzzles illustrate the issue. They represent examples of apparent *irrationality* in law which we believe are perfectly understandable in terms of common behaviour and are probably desirable.

Puzzle I:[3] A seller sells an automobile to a buyer (B1), permitting B1 to take possession of the car before payment in full. The seller reserves *ownership* of the car until full payment. Before full payment, B1 enters into an agreement with a sub-buyer (B2) who pays B1 a downpayment of $500 for the car, with delivery arranged for the following day. Before delivery to the sub-buyer the original seller discovers the sub-sale. In a dispute between the original seller claiming ownership of the car, and B2, who also claims ownership, the *original seller* will win.[4]

---

[1] O.W. Holmes, *The Common Law* (Boston: Little, Brown, 1881) at 94. As we point out later, it is not as simple as it might first appear to know *where* losses lie, in order to decide where they should fall. As well, it is clear to us that the idea to which Holmes is alluding has application far beyond *accidents*, and may inform our understanding of property rights, contract remedies and regulatory policy.

[2] The definition of efficiency which we employ in this paper is quite simple. A legal decision or rule is efficient if, after its application, at least one person is better off, and no one is worse off than before. We might adopt a Kaldor-Hicks definition of efficiency which permits hypothetical compensatory wealth transfers between the affected parties — whether we choose a Pareto or such a potential Pareto superiority criterion does not matter to our thesis.

[3] This puzzle is equally apparent in the treatment of sales by *sellers* who remain in possession of goods. Here, sales legislation which deals with sales by *sellers in possession*, generally expropriates the interest of the first buyer only where the second buyer has received delivery of the property.

[4] The original seller wins, since at common law the first buyer could not transfer any better *title* than she or he had. Thus the first sub-buyer could not receive any better ownership claim than *the person selling to this sub-buyer*, and would loose to the original owner. See *Cole* v. *North Western Bank* (1875), L.R. 10 C.P. 354 at 362, [1874-80] All E.R. Rep. 486, 44 L.J.C.P. 233 (Ex.D.) See

A seller sells an automobile to a buyer (B1), permitting B1 to take possession of the car before payment in full. The seller reserves *ownership* of the car until full payment. Before full payment, B1 enters into an agreement with a sub-buyer (B2) who pays B1 a downpayment of $500 for the car, with immediate delivery arranged from B1 to B2. After delivery to the sub-buyer the original seller discovers the sub-sale. In a dispute between the original seller claiming ownership of the car, and B2, who in this case has taken delivery of the car, and who also claims ownership, the *sub-buyer* wins![5]

Puzzle II: A seller contracts with a buyer to purchase a piece of jewellery. The buyer pays $100 as a deposit and agrees to pay $200 per month for 12 months to pay off the purchase price of $2,500. The buyer also agrees that on default, the seller will be able to retain any money which have been paid as of the date of default. The buyer defaults after 6 months, the seller repossesses the jewellery and retains the $1,300 which has been paid. If the buyer seeks to set aside the *forfeiture clause* and recover the amount he or she has paid, then the buyer will only be able to do so if it would be "unconscionable for the seller to retain the money."[6] In the vast-majority of cases, the courts will leave the money with the seller.

A seller contracts with a buyer to purchase a piece of jewellery. The buyer pays nothing down, and agrees to pay $200 per month for 12 months to pay off the purchase price of $2,400. The buyer agrees that

---

generally, M.D. Chalmers, *The Sale of Goods Act, 1889* (London: William Clowes & Sons, 1905) at 57-66.

[5] The outcome is produced through the operation of section 25(2) of the original English *Sale of Goods Act* (U.K.), 1893, c. 71, which re-enacted with some slight modifications, section 9 of the *Factors Act* (U.K.), 1889, c. 45 [hereinafter *Factors Act, 1889*]. However, that section only operates where the sub-buyer obtains *delivery of the goods or transfer of documents of title*, and thus does not protect mere contract or ownership expectations.

We should also point out that this example suggests that Holmes's insight is equally accurate of legislative action, at least where the legislative institution is purporting to create a framework for contract and ownership disputes analogous to the framework established at common law.

Finally, some insight into the reasons we protect the sub-buyer who has possession of the goods may be gained from early decisions which refused to extend the equivalent section of the *Factors Act, 1889*, to lessees, even where the lessee had entered into a lease-option arrangement. See *Helby* v. *Matthews*, [1895] A.C. 471, 64 L.J.Q.B. 465, 60 J.P. 20 (H.L.).

[6] See *Stockloser* v. *Johnson*, [1954] 1 Q.B. 476, [1954] 2 W.L.R. 439, [1954] 1 All E.R. 630. While there is no empirical data on this point, our impression, and the impression of most commentators, is that judicial discretion to take wealth away from the seller, and thus return monies to the buyer, is only very rarely exercised, compared to their willingness to review *penalty clauses*. See S.M. Waddams, *The Law of Contracts* (Toronto: Canada Law Book, 1977) at 277-81.

on default $1,300 will be paid to the seller as liquidated damages. The buyer defaults, and the seller repossesses the jewellery and sues to recover the $1,300 as liquidated damages. The liquidated damages clause will, however, be held unenforceable where the buyer persuades the court that the clause was *intended as a penalty*. In a large percentage of the cases, the courts will refuse to enforce the clause and leave the money with the buyer.[7]

The *puzzle* in these examples is that the outcome differs depending on the *location* of the object of the dispute. In Puzzle I, the issue of which claimant's ownership interest is expropriated depends on whether the sub-buyer has or has not taken delivery of the car. In Puzzle II, the issue of the enforceability of the private damage assessment term depends on whether the money has or has not been transferred to the seller.

Many writers have seen the similarity between the penalty and forfeiture cases and have argued forcefully for solutions which would *reconcile*[8] the two doctrines, or which would eliminate the "illogical distinction drawn by the existing law between penalty clauses and the doctrine of forfeiture."[9] It does not matter, of course, whether the proponents of reform would prefer the penalty/liquidated damages solution over the forfeiture clause solution,[10] or the converse. The point is that *rationality* to some reformers means that the situations should be treated the same way.

While the legal treatment of these cases appears anomalous, the outcomes reflect real and important differences in people's valuations of losses relative to gains, and in their judgments of fairness predicated on this disparity. The differences in valuations of gains and losses appear to go a long way towards explaining the *rationality* of the apparently illogical puzzles presented above, and of many others as well.[11] These puzzles can be explained as manifestations of the idea that *possession*[12] losses are much more important than foregone gains, and the law takes such real differences into account. In each case the legal rule treats economic gains and losses differently, even though nominally commensurate. These doctrines reflect powerful human sentiments:

---

[7] The distinction between the treatment of liquidated damages or penalties and forfeiture clauses becomes even more problematic where the defendant has promised to pay a deposit which is subject to forfeiture on breach, and then fails to pay it. The cases are split on this third situation. Some hold that the buyer must pay, that is, treating the arrangement as a forfeiture, while others hold that the arrangement represents a penalty clause. See *Hinton* v. *Sparkes* (1868), L.R. 3 C.P. 161, 37 L.J.C.P. 81, 17 L.T. 600 (buyer liable for promised amount, even though penal); *Dewar* v. *Mintoft*, [1912] 2 K.B. 373, 81 L.J.K.B. 885, 106 L.T. 763 (amount not recoverable by seller as a penalty). See S.M. Waddams, *The Law of Damages* (Toronto: Canada Law Book, 1983) at 540.

[8] See Waddams, *ibid.*

[9] Ontario Law Reform Commission, *Report on Sale of Goods* (Toronto: Minister of the Attorney General, 1979) at 425. Law Commission, *Penalty Clauses and Forfeiture of Monies Paid* (Working Paper No. 61) (London: Her Majesty's Stationery Office, 1975) at p. 51-54.

[10] The Law Commission of England would prefer that both be treated like penalty clauses are now treated, while Waddams would prefer that both be treated like forfeiture clauses. See Law Commission, *Ibid.*, and S.M. Waddams, *supra* note 7 at 540.

> bare expectations were less important than expectations allied to present rights, especially rights of property. Hume and Adam Smith, for example, both said that expectations arising out of rights of property deserved greater protection than expectations to something which had never been *possessed*. To deprive somebody of something which he merely expects to receive is a less serious wrong, deserving of less protection, than to deprive somebody of the expectation of continuing to hold something which he already *possesses*.[13] (emphasis added)

II. THE VALUATION DISPARITY

In each of the cases of judicial choices which we explore in this paper, greater weight is given to actual out of pocket losses than to the opportunity cost of foregone gains. Such asymmetries in the valuations of gains and losses are not at all consistent with the assurances of most economists that valuations of gains and losses are equivalent or with generally accepted principles of analysis based on such assumptions of people's preferences.

---

[11] Common law judges seemingly apply Aristotelian concepts of corrective justice by restoring parties to the position they were in prior to the wrongful act. Law as it has developed in most common law jurisdictions implicates ideas of corrective, rather than distributive justice. Corrective justice is a system for preserving what we have, for minimizing disruptions of feelings of loss, and not for bringing about new distributions of wealth.

Maintaining entitlements is implicit in law in three ways. First, compensation will usually be considered in terms of remedying actual losses to plaintiffs rather than in compensating for foregone gains. Second, and more important, law will represent itself as a passive political institution, demanding justification for transfers of wealth from defendants to plaintiffs. Finally, judicial discretion implicit in fact-finding, interpretation, and rule-selection and rule-application may incorporate concepts of fairness which reflect the idea that losses are much more important than foregone gains.

[12] While we use the word *possession* to denote the critical element in these cases as well as in the empirical studies which we describe, we are not suggesting that a physical connection or even legal entitlement is a prerequisite to the occurrence of the phenomenon. The reference position from which changes are perceived as gains or losses may well depend on factors other than physical possession.

[13] See P.S. Atiyah, *The Rise and Fall of Freedom of Contract* (Oxford: Clarendon Press, 1979) at 428.

There is no dispute that the economic value of both gains and losses is measured by what people are willing to sacrifice. That is, gains are valued by a payment measure, and losses are valued by a compensation measure. For example, Posner suggests, "[T]he economic value of something is how much someone is willing to pay for it or, if he [or she] has it already, how much money he [or she] demands to part with it."[14] Similarly, Michelman suggests, "[B]enefits are measured by the total number of dollars which prospective gainers would be willing to pay to secure adoption, and losses are measured by the total number of dollars which prospective losers would insist on as the price of agreeing to adoption."[15]

Although income effects, or limits on ability to pay, may cause the two measures to differ, these are normally not a factor of any practical importance and can safely be ignored in most applications. The common view is that, "for many goods, services, and amenities that command a modest fraction of the consumer's budget, the differences between [the] ... measures are trivial."[16] Consequently, the usual advice is that "as a practical matter it usually does not make much difference which of these two approaches ... is adopted."[17]

Common preferences and reactions to actual choices are, however, not consistent with the assertion of equivalence between the measures of loss and gain. Nor, it seems, are judicial reactions and choices.

The traditional assumption of equivalence between the willingness-to-pay and the compensation-demanded measures of value is largely an empirical assertion that is common to both economic and legal analysis. It is based on the behavioural assumption that people assess gains and losses by comparing how well off they would be with more or with less of something — by for example comparing how they would feel with their present wealth with how they would feel with their present holdings plus $100; or by comparing their level of welfare with an injury and their economic well-being without one.

---

[14] R. Posner, *An Economic Analysis of Law*, 3rd ed. (Boston: Little, Brown, 1986) at 11.

[15] F.I. Michelman, "Property, Utility, and Fairness: Comments on the Ethical Foundation of 'Just Compensation' Law" (1967) 80 Harv. Law Rev. 1165 at 1214.

[16] A. Randall, *Resource Economics* (New York: Wiley, 1987) at 244.

[17] S.E. Rhoads, *The Economist's View of the World* (Cambridge: Cambridge University Press, 1985) at 125.

New empirical evidence indicates that instead of comparing alternative end states, people's actual choices and actions suggest that they evaluate gains and losses in terms of changes from some reference position. This evidence also indicates that they value losses from this neutral point much more than they do gains beyond it: "[T]he aggravation that one experiences in losing a sum of money appears to be greater than the pleasure associated with gaining the same amount."[18]

Virtually all controlled evaluation tests, as well as the commonplace reactions of people to real choices, point to large and persistent differences between the valuation of losses and foregone gains and do not confirm the traditional assumption that these two measures of value are equal. Large differences were first noted in survey studies of people's valuations of various losses of environmental assets or the degradation of environmental quality. For example, a sample of duck hunters said they would be willing to pay an average of $247 to save a marsh area used by ducks, but would demand an average of $1044 to accept the identical loss.[19] Similarly, a survey of anglers yielded a payment value of $35 and a compensation value of $100 for the loss of a fishing area.[20] The disparity between the two measures of value was also reported to vary from $43 to $120 for the loss of a fishing pier;[21] from $22 to $93 for the loss of a local postal service;[22] from $21 to $101 for the loss of a goose hunting permit;[23] from $54 to $143 for the loss of an opportunity to hunt elk;[24] and from $40 to $833 for the loss of the to chance to hunt deer.[25] We are not aware of any surveys reporting equivalence between the two measures.

---

[18] D. Kahneman & A. Tversky, "Prospect Theory: An Analysis of Decision Under Risk" (1979) 47 Econometrica 263 at 279.

[19] J. Hammack & G.M. Brown, *Waterfowl and Wetlands: Toward Bioeconomic Analysis* (Washington: Resources for the Future, 1974).

[20] W.F. Sinclair, *The Economic and Social Impact of Kemano II Hydroelectric Project on British Columbia's Fisheries Resources* (Vancouver: Fisheries and Marine Service, Department of the Environment, 1976).

[21] N.D. Banford, J.L. Knetsch & G.A. Mauser, "Feasibility Judgements and Alternative Measures of Benefits and Costs" (1980) 11 J. of Bus. Admin. 25.

[22] *Ibid.*

Later studies based on real exchange experiments have provided more stringent tests than the earlier ones which were based on hypothetical survey questions. The results have been essentially the same. Even when exchanges of real goods and actual cash payments motivated the evaluations, the compensation demanded to give up an entitlement far exceeded the comparable payment measures of value. For example, people required about four times more money to give up a lottery ticket than they would be willing to pay to acquire one.[26] The values were reported to be $25 and $172 for the actual loss of a permit to hunt deer in the northern United States.[27]

A more recent series of real exchange experiments has affirmed the persistence of the evaluation disparities over repeated valuations; and has eliminated the possibility that the differences might be attributable to transaction costs or strategic behaviour on the part of participants.[28] The results indicate that, consistent with other evidence and contrary to conventional assertions, people value losses much more than gains. The differences are pervasive and large. They persist over repeated valuations and they are not the result of inhibitions posed by transaction costs.

An indication of the general findings is illustrated by the results of a simple exercise involving people's actual choices between a coffee mug and a large, four hundred gram chocolate bar.[29] Each member of one group of seventy-six participants was given a mug and told that it was theirs to keep and take with them. They were then given a chance to give up their mug to obtain a chocolate bar. A second group of eighty-seven participants were given the opposite choice — to give up chocolate bars to obtain coffee mugs.

---

[23] R.C. Bishop & T.A. Heberlein, "Measuring Values of Extra-Market Goods: Are Indirect Measures Biased?" (1979) 61 Am. J. Agric. Econ. 926.

[24] D. Brookshire, A. Randall & J. R. Stoll, "Valuing Increments and Decrements in Natural Resource Service Flows" (1980) 62 Am. J. Agric. Econ. 478.

[25] R.C. Bishop & T.A. Heberlein, "Does Contingent Valuation Work?" in R.G. Cummings, D.S. Brookshire & W.D. Schulze, eds., *Valuing Environmental Goods* (Totowa, N.J.: Roman & Allanheld, 1986) 123.

[26] J.L. Knetsch & J.A. Sinden, "Willingness to Pay and Compensation Demanded: Experimental Evidence of an Unexpected Disparity in Measures of Value" (1984) 99 Q. J. of Econ. 507.

[27] R.C. Bishop & T.A. Heberlein, *supra* note 25.

[28] D. Kahneman, J.L Knetsch & R.H. Thaler, "Experimental Test of the Endowment Effect and the Coase Theorem" (1990) 98 J. Pol. Econ. 1325.

Traditional economics predicts that, in the absence of any significant transaction costs, about the same proportion of participants in each of the groups will prefer mugs to chocolate bars, and will choose accordingly. The actual results were in sharp contrast to this prediction: 89 per cent indicated a preference for the mug when initially given a mug, and only 10 per cent revealed a similar preference for mugs when they had to give up a chocolate bar to obtain one. The relative value of mugs and chocolate bars varied greatly and depended on whether the evaluation was made in terms of a gain or a loss. The mug was valued more when it had to be given up to obtain a chocolate bar; and was valued less when the chocolate bar had to be given up to obtain a mug. The influence of income constraints and wealth or income effects were entirely eliminated in the exercise leaving the choices dependent on preferences of the individuals. These choices showed that gains and losses were valued very differently.

A further test for the equivalence or non-equivalence of valuations of gains and losses was modeled on the Coase Theorem — a mainstay of the economic analysis of law and the basis for many legal policy prescriptions. A major conclusion of the Coase proposition is that, in the absence of transaction costs and wealth effects, people are presumed to make mutually advantageous exchanges to ensure that resources are put to their most valued use.[30] "Since a receipt foregone of a given amount is the equivalent of a payment of the same amount,"[31] final allocations of entitlements are assumed to be independent of initial entitlements. "[T]he ultimate result (which maximizes the value of production) is independent of the legal position if the pricing system is assumed to work without cost."[32]

---

[29] J.L. Knetsch, "The Endowment Effect and Evidence of Non-Reversible Indifference Curves" (1989) 79 Am. Econ. Rev. 1277.

[30] R.H. Coase, "The Problem of Social Cost" (1960) 3 J. L. & Econ. 1.

[31] *Ibid.* at 7.

The Coase Theorem test was conducted by randomly giving one member of each of a large number of paired participants a good — a large chocolate bar — which they were told was theirs to take home or to sell if the person with whom they were paired made a sufficiently attractive offer. The experiment was deliberately arranged so that the potential buyers had been given a larger windfall sum of money than the sellers. Despite this effort to encourage buyers to use their unanticipated gain to make larger offers to acquire the goods, few transactions were concluded.[33] In spite of the entitlements' random distribution, the people holding them demanded much more to give them up than potential buyers were willing to pay to acquire them. The valuations were *not*, contrary to the Coase prediction, independent of the initial assignments of entitlements.[34]

Individuals have now repeatedly been shown to exhibit disparities between gain and loss valuations in experimental settings, as reported by many investigators using a variety of methods to evaluate widely varied assets. As well, people's actual behaviour in making everyday choices increasingly has been observed to be consistent with these findings. For example, the valuation disparity, and consequent reluctance to sell at a loss, is observed in the greater volume of house sales when prices are rising, over the number when they are falling. This can also be observed in the smaller volume of sales of securities that have declined in price relative to those for which prices have increased.[35] Firms frequently are reluctant to divest themselves of plants and product lines even though they would not consider buying these same assets, and stock prices often rise when they do give them up.

A further illustration of the differing valuations of gains and losses is provided by responses to recent automobile insurance legislation in two American states. In both jurisdictions people are given a choice between cheaper policies that limit rights to subsequent recovery of further damages and a more expensive policy permitting such actions. Importantly, the default option differs: the *reduced* rights policy is offered in New Jersey unless it is given up; and *full* rights policy is given in Pennsylvania unless the less expensive option is specified. Given the minimal costs in both states of choosing either option, and the large amounts of money at issue, the results have been dramatic. At last count over 70 per cent of New Jersey automobile owners have adopted the reduced rights policy, but fewer than 25 per cent of Pennsylvanians have done so.[36]

---

[32] *Ibid.* at 8.

[33] *supra* note 28.

[34] The major efficiency implications of the valuations of gains and losses have been demonstrated in H. Hovenkamp, "Legal Policy and the Endowment Effect" (1991) 20 J. Legal Stud. 225 at 230.

[35] H. Shefrin & M. Statman, "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence" (1985) 40 J. Fin. 777.

The differing weights attached to gains and losses have also been found to influence people's judgments of what they regard as acceptable or unfair behaviour in economic relationships. If an action is seen to impose a loss on one party for the benefit of another, this will nearly always be widely seen to be unfair — quite apart from whatever economic justification might exist. For example, raising prices in response to sudden shifts in demand is seen to benefit the seller, who has not incurred any cost increase, at the expense of the buyer and is therefore judged to be unfair. Similarly, cutting wages when unemployment increases is thought to be unfair by the vast majority of people because the employer benefits in direct proportion to the worker's loss. Reducing a customary bonus payment to workers is apparently viewed as a less aversive foregoing of a gain and is therefore far more acceptable than an equivalent reduction in wages, which is commonly seen as imposing a loss on workers.[37]

Raising rents to a new as opposed to a sitting tenant, or cutting wages to a new rather than an old employee are generally considered to be fair because the benefit to the landlord or the employer is not seen to be gained at the expense of a loss to the other party. The evidence also suggests a willingness of most people to back up their judgments with sacrifices to punish unfair behaviour and reward what they take to be fair dealings.[38] A major motivation for these judgments appears to be that losses matter more to people than do foregone gains.

---

[36] J. Meszarose *et al.*, *Framing, Loss Aversion and Insurance Decisions* (Working Paper) (Philadelphia: University of Pennsylvania Press, 1991).

[37] D. Kahneman, J.L. Knetsch & R.H. Thaler, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market" (1986) 76 Am. Econ. Rev. 728.

This strong intuition to value losses more than commensurate gains, and the implications for the resolution of competing claims, was summarized earlier by Holmes in the following terms:

> It is in the nature of man's mind. A thing which you have enjoyed and used as your own for a long time, whether property or an opinion, takes root in your being and cannot be torn away without your resenting the act and trying to defend yourself, however you came by it. The law can ask no better justification than the deepest instincts of man.[39]

While there has been very little empirical work investigating this phenomenon among legal actors, Stewart Macaulay's seminal work in the sociology of law almost three decades ago, offers some support for the appearance of this phenomenon among members of at least one business community when they confront legal disputes.[40] Macaulay interviewed sixty-eight businesspeople and lawyers representing forty-eight companies and six law firms. All but two of the companies were engaged in manufacturing with plants in Wisconsin. In his investigation of dispute settlements, Macaulay began with the formal assumption that, on breach of contract, the breaching party, say the buyer, would be legally obligated to pay all of the seller's wasted expenses up to the time of breach, plus anticipated lost profits. However, the responses of the purchasing agents and sales personnel revealed attitudes consistent with the experimental data discussed above. That is, they uniformly believed that all they ought to recover or to pay in damages was an amount representing the seller's actual expenses:

> [A]ll ten of the purchasing agents [that were] asked about cancellation of orders once placed indicated that they expected to be able to cancel orders freely subject to only an obligation to pay for the seller's major expenses such as scrapped steel. All seventeen

---

[38] D. Kahneman, J.L. Knetsch & R.H. Thaler, "Fairness and the Assumptions of Economics" (1986) 59 J. Bus. 285.

[39] O.W. Holmes, "The Path of Law" (1897) 10 Harv. L. Rev. 457 at 477. The only difference revealed by the empirical evidence is that the reluctance to give up something does not necessarily only occur after "a long time," but may well set in immediately — once a reference position is perceived which leads to a change being viewed as a loss, then the valuation will respond accordingly. See, particularly, *supra* note 28.

[40] See S. Macauley, "Non-Contractual Relations in Business: A Preliminary Study" (1963) 28 Am. Soc. Rev. 55.

> sales personnel asked reported that they often had to accept cancellation. One said "You can't ask a man to eat paper [the firm's product] when he has no use for it."[41]

The empirical findings discussed earlier, together with Macauley's evidence, suggest that people value gains and losses from some neutral reference point or level. Losses from this reference are commonly weighed more than gains beyond it. The reference position therefore determines whether an adverse change is regarded as a loss or as a foregone gain, and whether a positive change is treated as a gain or as a reduction of a loss. Given the large disparities between valuations, these differences can have significant practical impacts.

III. THE DISPARITY AND JUDICIAL CHOICE

The legitimacy of the common law has largely been based on its decentralized, *ad hoc*, incremental development by judges who have little enforcement power, and who therefore respond to an intuitive, non-empirical interpretation of community mores, and individual preferences. If such interpretations underlie individual judgments, then a phenomenon as pervasive as the valuation disparity would be expected to be implicated, either explicitly or implicitly, in the development of legal doctrine. Indeed, as expected, the idea that *losses count more than expected gains* is one which is encountered in an enormous range of ideas in law.[42]

An important way in which the valuation disparity has been incorporated in law is through the widespread recognition of possession as a foundation for the declaration or recognition of legal entitlements.[43]

---

[41] *Ibid.* at 64.

[42] Other, independent, *ad hoc* explanations can be offered for many specific results, but the valuation disparity alone appears uniquely robust as a consistent explanation across a wide range of apparently independent legal issues.

[43] One of the most articulate writers on the relationship of possession and the human condition is C.B. MacPherson who, in *The Political Theory of Possessive Individualism* (Oxford: Clarendon Press, 1962) at 3, wrote that "[t]he human essence is freedom from dependence on the wills of others, and freedom is a function of possession." See also, British Columbia Law Reform Commission, *Wrongful Interference with Goods* (Working Paper No. 67) (Vancouver: British Columbia Law Reform Commission, 1992) at 21.

The often misused expression that "possession is nine-tenths of the law"[44] is a manifestation of a judicial intuition that possession of things is intimately connected to establishing the endowment effect and underlies judicial choices favouring protection against losses over foregone gains.

Both legal writers and the designers of the empirical studies discussed earlier have used the term possession in extremely ambiguous ways. But regardless of one's choice of definition, it is incontrovertible that judicial sensitivity to the notion of possession, or seisin, in English law is deeply rooted and utterly pervasive: "In the history of our law there is no idea more cardinal than that of seisin."[45]

Frederick Pollock,[46] while noting that it was difficult to obtain a consistent doctrine or consistent terminology,[47] offered several theories on why judge-made common law presumes in favour of possession.[48] Utilitarian justifications for the protection of possession range from a concern with reducing the risk of civil disobedience — a response to the risk of personal injuries associated with self-help dispossession — to the protection of ownership interests. But once again, we find references to the *natural* expectation that a person should and will be allowed to keep what he or she possesses until someone has proved a better title.[49]

---

[44] Historians have traced its development from Roman rather than Germanic Law. See S.S. Peloubet, *A Collection of Legal Maxims in Law and Equity, with English Translations* (Littleton, Colo.: F.B. Rothman, 1985) at 225; P. Hackerston, *A Collection of Latin Maxims and Rules* (1823) at 127; Sir F. Pollock & F.W. Maitland, *The History of English Law: Before the Time of Edward I*, vol. 2, 2d ed. (Cambridge: Cambridge University Press, 1968) at 29. In *The Corporation of Kingston-upon-Hull* v. *Horner*, [1774] Lofft. 592, 98 E.R. 807 at 815, Lord Mansfield C.J. said that "Possession is very strong; rather more than nine points of the law." He used the expression in a case concerning prescriptive rights where, in the absence of evidence of a Crown Grant, one was presumed to exist, thus supporting a possessory title. By 1881 in *Beddal* v. *Maitland* (1881), 17 Ch.D. 174 at 183, Sir Edward Fry was noting that the old saying that possession is nine points of the law was created by a forcible entry statute which resulted in a man in possession being able to use force to keep out a trespasser; however, if a trespasser had gained possession, the rightful owner could not use force to put him out, but had to appeal to the law for assistance. As late as 1946, *Wharton's Law Lexicon* was expressing the idea as possession constituting "nine points" of the law: J.J.S. Wharton, *Wharton's Law Lexicon*, 13th ed. by A.S. Oppe (London: Stevens, 1938) at 666. H.C. Black, ed., *Black's Law Dictionary*, 6th ed. (St. Paul, Minn.: West Publishing, 1990) at 1164 currently uses the phrase: "Possession is nine-tenths of the law."

[45] Pollock & Maitland, *ibid.* at 29. See also F. Pollock & R.S. Wright, *An Essay on Possession in the Common Law* (Oxford: Clarendon Press, 1888) at 1.

[46] Both with R.S. Wright (see Pollock & Wright, *ibid.*) and with F.W. Maitland (see Pollock & Maitland, *ibid.*).

[47] Pollock & Maitland, *ibid.* at 44. Pollock and Maitland make the point that "... so far as concerns our own English law we make no doubt that at different times and in different measures every conceivable reason for protecting possession has been felt as a weighty argument and has had its influence on rights and remedies," *ibid.* at 44.

[48] *Ibid.* at 40.

A. *Adverse possession*

The rule of adverse possession provides a means by which the user of property can successfully assert a claim of ownership over a prior owner. Many of the cases involve mistakes, for example, on the location of boundaries.[50] Others, however, reward those who make effective use of an asset over an owner who has effectively abandoned it.[51]

The efficiency advantages of awarding titles to adverse possessors usually include the reduction of administrative costs of establishing rightful ownership, and the encouragement to make productive use of assets which are left unused by their owners.[52] Posner suggests a further efficiency advantage of the rule of adverse possession, turning on possible differences between the valuations of owners and possessors because of differences in the marginal value of their wealth levels:

> The adverse possessor would experience the deprivation of property as a diminution in his wealth; the original owner would experience the restoration of the property as an increase in *his* wealth. If they have the same wealth, then probably their combined utility will be greater if the adverse possessor is allowed to keep the property.[53]

Posner chooses to interpret Holmes's views[54] of the disparity between feelings of gains and losses as "a point about diminishing marginal utility of income."[55] The evidence discussed earlier suggests instead that Holmes's notion is a prescient articulation of the endowment effect — that is, the disparity between the valuation of gains and losses. The reference positions are likely such that taking away from a current user would likely be valued as a loss, and giving back to the original owner would likely be valued as a gain. And given the greater valuations of losses over gains, the rule seems consistent with maximizing joint welfare, quite apart from any concern with the diminishing marginal utility of income or any incentives to encourage investment and use.

---

[49] *Ibid.* at 41-43.

[50] Distinctions in granting variances are commonly made between not requiring destruction of an improvement built too close to a neighbour by mistake and proposing to do so, and between wilfully ignoring a zoning ordinance and doing so by mistake.

[51] While now largely matters of statute in most jurisdiction, the rules grew out of resolutions favouring current users because of the difficulty of establishing old entitlements due to lost records and fading memories. See R. Cooter & T. Ulen, *Law and Economics* (Glenview, Ill.: Scott, Foresman & Company, 1988).

[52] See *Ibid.*; W.Z. Hirsch, *Law and Economics: An Introductory Analysis* (New York: Academic Press, 1979).

[53] R. Posner, *An Economic Analysis of Law*, 3d ed., (Boston: Little Brown, 1986) at 70.

[54] See Holmes, *supra* note 39 at 477.

The potential gains depend on how well the limitations and restrictions on length of time and behaviour needed to acquire title reflect reference positions of the parties and their consequent views of gains and losses. If a person makes very temporary use of a parcel of land while the owner is absent for a short time, the reference position is unlikely to shift and an award of title would then be taken as a gain. If the use continues over many years, a reasonable expectation of continued use would probably develop so that any unexpected termination would be viewed as a more important loss. An owner deprived of title after a short absence would surely take this as a loss, and the discovery that ownership extended several metres beyond an old boundary fence would likely be regarded as a less valued gain.

The rules of adverse possession may be consistent with lowering administrative costs and encouraging use. However, neither addresses the disparate valuations — and consequent efficiency outcomes — that depend on perceptions of gains and losses, and the specification of requirements in terms of reference positions. A short limitation period might well induce use of *unused* property, but if such a specification leaves the original owner in the domain of losses and the possessor feeling that the title is a gain, the result is likely to be an inefficient change in which the gain is outweighed by the loss.

B. *Limitations on Recovery of Lost Profits*

A second, more specific example of the operation of the valuation disparity on judicial decisions is the reluctance of judges to compensate for *lost profits*, whether the claim is framed in contract or tort. As Adam Smith wrote:

---

[55] *supra* note 53 at 70. This was previously observed by R.C. Ellickson, "Bringing Culture and Human Frailty to Rational Actors: A Critique of Classical Law and Economics", (1989) 65 Chicago-Kent Law Review 23 at 38.

> [W]e naturally depend more on what we possess than what is in the hands of others. A man robbed of five pounds thinks himself much more injured than if he had lost five pounds by a contract.[56]

When permitting the recovery of economic losses in tort law, the idea is reflected in the consistent distinction which judges draw between "loss by way of expenditure and failure to make gain."[57] While recovery of expenditures is sometimes permitted, recovery of foregone gains is not.[58] Several recent decisions have explicitly recognized the distinction. In one, *Dominion Tape of Canada Ltd.* v. *L.R. McDonald & Sons Ltd.*,[59] several bales fell from a trailer, and hit a hydro pole cutting power to the plaintiff's plant. The plaintiff successfully sued to recover wages paid to employees which represented *positive outlays*; but could not recover loss of profits, which were only *negative losses*, consisting of a "mere deprivation of an opportunity to earn an income." The latter, while foreseeable, was judged too remote.[60]

---

[56] A. Smith, *Lectures on Justice, Police, Revenue and Arms*, ed. by E. Cannan (Oxford: Clarendon Press, 1896) at 131.

[57] Thus in J.F. Clerk, *Clerk & Lindsell on Torts*, 15th ed. by R.W.M. Dias (Agincourt, Ont: Carswell, 1982) at 32-33, 371-385 the authors typically offer the formal statement that recovery has been extended to economic losses representing expenditures, but "the general rule is that loss of profit *per se* is not actionable: there is a no duty-situation." Similarly, in *Weller & Co.* v. *Foot and Mouth Disease Research Institute* (1965), [1966] 1 Q.B. 569, [1965] 3 All E.R. 560, the Court held that the loss of profit did not constitute a harm of a sort which the law would remedy, and that the rule against recovery is independent both of negligence and foreseeability.

[58] The House of Lords decision in *Junior Books Ltd.* v. *Veitchi* (1982), [1983] 1 A.C. 520, [1982] 3 W.L.R. 477 (H.L.) [hereinafter *Junior Books*], which might have presented a radical development in compensation law, and which permitted recovery of lost profits in a non-contractual setting, has not been widely followed. See J.F. Clerk, *Fifth Cumulative Supplement to the Fifteenth Edition*, by R.W.M. Dias (London: Sweet & Maxwell, 1987) at para. 10-14 which describes six limiting factors on the *Junior Books* decision.

[59] [1971] 3 O.R. 627, 21 D.L.R. (3d) 299 (Co. Ct.). See also *MacMillan Bloedel Ltd.* v. *Foundation Co. of Canada Ltd.* (1977), 75 D.L.R. (3d) 294, [1977] 2 W.W.R. 717, 1 C.C.L.T. 358 (B.C.S.C.) (employer may be permitted to recover actual wages paid to employees); *Ontario (A.G.)* v. *Crompton* (1976), 14 O.R. (2d) 659, 74 D.L.R. (3d) 345, 1 C.C.L.T. 81 (H.C.J.) (specific expenditure incurred in putting out fire recoverable).

[60] This approach received some support from Wilson J.A. as an innovative approach in *Ontario (A.G.)* v. *Fatehi* (1981), 34 O.R. (2d) 129 at 140, 127 D.L.R. (3d) 603 at 615 (C.A.), rev'd [1984] 2 S.C.R. 536, 15 D.L.R. (4th) 132.

While lost profits are said to be recoverable in contract law, it is difficult to defend the proposition that the *expectation* interest of a non-breaching party is recognized in contract actions in the same fashion as are *actual* losses.[61] As Atiyah argued, the costs of depriving a breaching party of his or her wealth are considered by most to outweigh the expected benefit to the nonbreaching party:

> [I]t might well be thought by most people that the inconvenience to the promisor of being held to his contract would be enough to outweigh the *prima facie* desirability of not disappointing the promisee.[62]

Fuller and Perdue in their seminal article, "The Reliance Interest in Contract Damages,"[63] argued that the public interest in redressing wrongs varies directly with the type of claim which the complainant makes. In their eyes, it is relatively easy to justify claims to compensation where the plaintiff is seeking return of wealth transferred to the defendant, and almost as easy to justify claims to compensation for actual out of pocket expenses incurred as a result of the breach of contract. However, while contract law protects expectations of gain:

> [I]n passing from compensation for change of position to compensation for loss of expectancy ... [t]he law no longer seeks ... to heal a disturbed *status quo*, but ... assumes a more active role. *With the transition, the justification for legal relief loses its self-evident quality.*[64] [emphasis is mine]

Fuller and Purdue justify the protection of economic expectations by explaining that the promisee's assets are *actually* reduced but only in the sense that this would be understood "according to the modes of thought which enter into our economic system."[65] Moreover, they explain that the reason for the protection of expectations is that the, adoption of a rule which protects foregone opportunities is, in effect, the most effective means of encouraging reliance and thus compensating for actual incurred losses!

---

[61] Thus Hugh Collins writes that "In practice ... the courts rarely countenance the ... head of damages for anticipated profits." H. Collins, *The Law of Contract* (London: Weidenfeld & Nicolson, 1986) at 181-82. See also, L.E. Wolcher, "The Accommodation of Regret in Contract Remedies" (1988) 73 Iowa L. Rev. 797 at 873.

The most vocal proponent of the thesis that contract liability is based on reliance and restoration of benefits, and thus that contract recovery should be similarly defined is Patrick Atiyah who has made this point on several occasions. See P.S. Atiyah, *The Rise and Fall of Freedom of Contract* (Oxford: Clarendon Press, 1979) at 763-779. See also, P.S. Atiyah, "Contracts, Promises and the Law of Obligations" (1978) 94 L.Q.R. 193.

[62] "Contracts, Promises and the Law of Obligations," *ibid.* at 216. Professor Scott has more pointedly written of the intuition that "a loss of $100.00 is more unpleasant than a gain of $100.00 is attractive." R.E. Scott, "Error and Rationality in Individual Decisionmaking: An Essay on the Relationship Between Cognitive Illusions and the Management of Choices" (1986) 59 S. Cal. L. Rev. 329 at 335.

[63] (1936) 46 Yale L.J. 52 at 53-57.

[64] *Ibid.* at 56-57.

In recent years, several judges have reinterpreted the application of *remoteness* limitations on contract damage recovery. Remoteness concepts are employed to place limits on damage awards by requiring that the risk of loss be "foreseeable with some degree of probability" at the time of contracting.[66] The doctrinal rules support the view that the limiting concepts are not the same in tort as in contract, and that tort law — employed to compensate for actual losses rather than forgone gains — "imposes a much wider liability."[67]

The valuation disparity has been explicitly recognized in *H. Parsons (Livestock) Ltd.* v. *Uttley Ingham & Co.*,[68] where the Court considered the recoverability of losses associated with the sale and installation of a defective feed hopper. Lord Denning reinterpreted earlier remoteness cases as establishing two distinct rules which could be understood as applying *not* to the doctrinal categories of tort and contract, but rather to the kinds of losses for which compensation was being sought. In the case of physical damage and actual expenses, the injured party should be able to recover losses which are foreseeable as resulting from a breach of contract even if, at the time of contract, there was only an extremely small probability of the loss occurring. However, claims representing loss of profit or loss of opportunities for gain are recoverable only if the risk of loss was contemplated as a serious possibility or a real danger.[69] The obvious result, whether or not one supports Lord Denning's re-articulation of the traditional remoteness rules,[70] is that the likelihood of recovery of losses is far greater than the likelihood of recovery of unrealized gains.

---

[65] *Ibid.*

[66] The common law rules defining remoteness concepts in contract and tort are expressed in *Koufos* v. *C. Czarnikow Ltd.* (1967), [1969] 1 A.C. 350, [1967] 3 W.L.R. 1491, [1967] 3 All E.R. 686 (H.L.), and *Overseas Tankship (U.K.) Ltd.* v. *The Miller Steamship Company* (1966), [1967] 1 A.C. 617, 3 W.L.R. 498 [1966] 2 All E.R. 709 (P.C.) (hereinafter *The Wagon Mound No. 2*), respectively. In the former case, all of the judges attempted to delineate the particular degree of probability with which the parties ought to have foreseen the risk of loss resulting from the breach of contract. *The Wagon Mound No. 2* served similar purposes in the law of torts — the issue being whether the defendant ought to be liable for all reasonably foreseeable risks regardless of their degree of probability, or whether some threshold statistical level of probability would trigger liability.

[67] See Lord Reid in *The Wagon Mound No. 2*, *ibid.* at 634.

[68] (1977), [1978] Q.B. 791, [1977] 3 W.L.R. 990, [1978] 1 All E.R. 525 (C.A.).

The principle reflected in this decision — that physical injuries and property damage should be given greater legal protection than unrealized expected economic gains — is again consistent with the empirical observation that people value losses more than foregone gains.

C. *Contract Modifications*

A third area where the valuation disparity manifests itself is in the distinctive treatment of performed and unperformed *intra-contractual promises*. For example, where a person promises to pay an additional amount of money for a previously arranged contractual performance,[71] the courts have uniformly denied the promisee the right to enforce the promise.[72] Thus, if a construction company agrees to pay an additional $20,000 to ensure delivery of steel which has already been contracted for, the court will not enforce the promise to pay the additional sum. However, if the construction company agrees to pay the additional $20,000 and actually transfers the money to the steel supplier, the promise is presumptively enforceable. The court will not order the money to be returned unless the construction company can demonstrate that the money was extorted from it under conditions of economic duress.[73]

---

[69] This idea was articulated by H.L.A. Hart & A.M. Honore, in *Causation in the Law* (Oxford: Clarendon Press, 1959) at 281-87. It is also explicitly recognized in some Civilian jurisdictions. For example, the German Civil Code limits contract recovery in certain cases to the "negative" or reliance interest, as discussed in G.E. Clos, *Comparative Law* (Littleton, Colo.: Fred B. Rothman & Co., 1979) at 226-236. Furthermore, Article 252 explicitly limits recovery of losses of particular gains to situations where the risk is foreseeable as a probable consequence of the breach. See Germany, *The German Civil Code*, trans. I.S. Forrester, S.L. Goren & H.-M. Ilgen (South Hackensack, N.J.: Fred B. Rothman & Co., 1975).

[70] The judgment is important because he confesses that the physical injury-property damage versus economic loss distinction operates *de facto* whatever formal test one accepts. See *supra* note 68 at 533 where he emphasizes a particular set of product liability cases in which personal injury claims were compensated "even though [manufacturers and retailers] had not the faintest suspicion of any trouble."

[71] The same anomaly operates where a creditor promises to accept less than the full amount owing from a debtor. At common law, the creditor could retract the promise even where the debtor had paid the money. See *Foakes* v. *Beer* (1884), 9 App. Cas. 605 (H.L.). However, the *Mercantile Law Amendment Act*, R.S.O. 1980, c. 265, s. 16 provides that *part performance* of the promise will extinguish the obligation. Again, the distinction between the two cases is simply where the money is *located*.

[72] See *Gilbert Steel Ltd.* v. *University Construction Ltd.* (1976), 12 O.R. (2d) 19, 67 D.L.R. (3d) 606 (C.A.).

The situations are indistinguishable except, again, for the location of the money. In the first example, the court will not order the transfer of wealth from the promisor to the promisee, preferring the *status quo*. In the second example, the court will not order the return of the wealth from the promisee to the promisor, again preferring the *status quo*. There are very persuasive reasons for singling out these *intra-contractual* promises for special treatment. They represent very real risks of extortion generated by reliance-based situational monopolies, introduce uncertainty relating to the authority to make the alleged modification of the agreement, and generate additional transaction costs to contractual dispute resolution. But those reasons apply with equal force to performed and unperformed promises. The distinction developed in the cases is, once more, consistent with the empirical data which points to significant welfare losses in the case of coerced transfers of wealth.

D. *Gratuitous Promises*

Judges have traditionally and consistently drawn a distinction between *giving* a gift to someone and *promising* to do so.[74] Performed gifts are enforceable, unperformed promises to give gifts are not. The undelivered gift cases are consistent with contract doctrines; the courts have not looked kindly on gratuitous promises[75] either at common law or at equity, and normally the rule is that they are not enforceable.[76] The major exception to the bargain theory of consideration is the enforcement of promises which have generated reasonable, detrimental reliance.[77]

---

[73] See R. Goff & G. Jones, *The Law of Restitution*, 2d ed. (London: Sweet & Maxwell, 1978). See *Peter Kiewet Sons' Co. of Canada* v. *Eakins Construction Ltd.*, [1960] S.C.R. 361, 22 D.L.R. (2d) 465; *Re Municipal Spraying & Contracting Ltd.* (1980) 15 B.L.R. 37, (*sub nom. Municipal Spraying and Contracting Ltd.* v. *Nfld.*) 153 A.P.R. 91 (Nfld. T.D.) (plaintiff entitled to recover $195,000 worth of additional work performed for defendant on ground that performance was rendered in response to threatened legal action and was not voluntary).

[74] See M.A. Eisenberg, "The Principles of Consideration" (1982) 67 Corn. L. Rev. 640; M.A. Eisenberg, "Donative Promises" (1979) 47 U. Chi. L. Rev. 1; *Halsbury's Laws of England*, 4th ed., vol. 20 (London: Butterworths, 1978) para. 62 at 36:

> where a gift rests merely in promise, whether written or verbal ... it is incomplete ... and the court will not compel the intending ... to complete and perfect it.

Eisenberg classifies donative promises into three groups: those which are informal and unrelied-upon; those which are formal but unrelied-upon; and those which are relied-upon.[78] Judges have consistently refused to enforce those gift promises which fall into the first and second categories. Several underlying rationales for the non-enforcement have been examined by several authors and various explanations have been offered. For example, Posner argues that the rule is perhaps reflecting an empirical *hunch* that "gratuitous promises tend to ... [involve] small stakes"[79] and therefore, that the social costs of enforcing the promise will generally exceed the utility of doing so. Swan and Reiter suggest that we are "naturally suspect" of gift promises. Because the law has provided rules for the completion of legally enforceable gifts, we should ask ourselves why those rules were not followed and whether, in fact, the donor really intended to complete the gift.[80]

An alternative explanation points to the small welfare loss associated with not receiving a gift, which is measured by the donee's foregone gain of the opportunity, relative to the welfare loss associated with depriving an existing owner of property of an entitlement, which is measured by the owner's asking price to give it up. This explanation, unlike the others, is consistent with the empirical data described earlier. Thus, Eisenberg argues that "lost expectation — a special form of disappointment — is among the least intense of injuries" and, therefore, is not worthy of legal remedy.[81] Since "lost expectation" is a foregone gain, Eisenberg clearly reveals an intuitive awareness of the valuation disparity and its role in understanding and predicting judicial choice. An unperformed gift promise is not treated as a loss. The "wrench of delivery"[82] has not been experienced by the promisor and no *real* loss has been experienced by the promisee.

---

[75] Gratuitous or donative promises are those which are given without the exchange of something of value, whether it be tangible property or another promise. *Currie* v. *Mesa* (1875), L.R. 10 Ex. 153, 23 W.R. 450; *Spruce Grove* v. *Yellowhead Regional Library Board* (1982), 44 A.R. 48, 143 D.L.R. (3d) 188, 21 M.P.L.R. 62 (C.A.).

[76] See for example A.G. Guest, *Anson's Law of Contract*, 26th ed. (Oxford: Clarendon Press, 1984) at 82.

[77] R.A. Posner, "Gratuitous Promises in Economics and Law" (1977) 6 J. Leg. Stud. 411; C. Fried, *Contract as Promise* (Cambridge: Harvard University Press, 1981) at 56.

[78] Eisenberg, "Donative Promises," *supra* note 74 at 6-11.

[79] See Posner, *supra* note 77 at 417. Posner argues that "Promises should not be enforced where the enforcement cost — to the extent not borne by the promisor — exceeds the gains from enforcement," and the only reason for enforcement would be an increase in net social welfare, at 414. It would, as Posner puts it at 417, be uneconomical to enforce casual social promises where the increment in utility to the promisor would be small if the promise were enforceable and the legal error costs high.

[80] See J. Swan & B.J. Reiter, *Contracts, Cases, Notes and Materials*, 3d ed. (Toronto: Emond Montgomery, 1985) at 223; M.D. Bayles, "Legally Enforceable Commitments" (1985) 4 L. & Phil. 312 at 338.

A more realistic assessment of gains and losses, taking the valuation disparity into account would predict that relied upon gratuitous promises will often be enforced because in those cases a *real* loss would have been incurred.[83] In *Skidmore* v. *Bradford*[84] a nephew was promised a gift of a warehouse by his uncle who, in addition to paying 1,000 pounds towards the purchase price of 5,000 pounds, had asked the owner to amend the agreement by writing in the nephew's name, and to make out the receipt in the nephew's name. Before he died, the uncle paid a further 500 pounds on the purchase price. Since the nephew had signed the agreement to purchase he was held liable for it. He paid, sued the uncle's estate to enforce the promise to give the warehouse as a gift, and *won*. There, the Court held that the nephew had "incurred that liability on the faith of the representations of the testator that he would give him the warehouse ..."[85]

Article 90(1) of the *Restatement (Second) of Contracts*[86] represents the modern version of this position. It provides that promises which can reasonably be expected to induce reliance are binding and should be enforced through remedies "limited as justice requires." While the cases discussing the appropriate remedy under Article 90 are not uniform, a substantial percentage award remedies limited by the promisee's reliance losses rather than expected benefits.[87]

---

[81] Eisenberg, "Donative Promises," *supra* note 74 at 3. See also, Eisenberg, "Principles of Consideration" *supra* note 74 at 656 where he says that "[t]his principle can be justified on several grounds, the most important of which is the low level of injury resulting from breach."

[82] P. Mecham, "The Requirement of Delivery in Gifts of Chattels and of Choses in Action Evidenced by Commercial Instruments" (1926) 21 Ill. L. Rev. 341 at 348-49.

[83] This is, in effect, the argument presented by P. Atiyah, *Consideration in Contracts: A Fundamental Restatement* (Canberra: Australian National University Press, 1971).

[84] (1869), L.R. 8 Eq. 134.

[85] *Ibid.* at 137.

[86] *Restatement (Second) of the Law of Contracts* § 90 (1981).

E. *Opportunistic Conduct*

Opportunistic conduct during contract formation and performance may take several forms including withholding information during contract formation, taking advantage of ambiguities in language, demanding performance in unintended situations, and failing to perform in order to take advantage of unexpected opportunities. Yet contract negotiations and performance will likely take place more effectively if trust is present and is generated by the process.[88] Risks of opportunism can be reduced by questioning the other's motives, honesty and future plans,[89] but it is inappropriate and probably dysfunctional to do so.

The treatment by the courts of one form of opportunistic behaviour — deliberate contract breaches motivated by attempts to capture unanticipated profitable opportunities — represents another example of the incorporation of the valuation disparity concept into judicial decision making. Here, the valuation disparity manifests itself in

---

[87] See *supra* note 63 at 64-5. See, for example, *Goodman* v. *Dicker*, 169 F.2d 684 (D.C. Cir. 1948); *Associated Tabulating Services* v. *Olympic Life Insurance Co.*, 414 F.2d 1306 (5th Cir. 1969).

[88] See R. McKean, "Economics of Trust, Altruism, and Corporate Responsibility" in E.S. Phelps, ed., *Altruism, Morality and Economic Theory* (New York: Russell Sage Foundation, 1975) 29; D. Collard, *Altruism and Economy: A Study in Non-selfish Economics* (Oxford: Martin Robertson, 1978) at 12; I.R. Macneil, *The New Social Contract* (New Haven: Yale University Press, 1980).

By trust we mean attitudes and behaviour which indicate that each person is willing to rely on the other to act fairly, and to take into account the other's welfare. Ian Macneil suggests that he uses solidarity and trust as equivalents, referring to a belief in future harmonious affirmative cooperation. , See I.R. Macneil, "Exchange Revisited: Individual Utility and Social Solidarity" (1986) 96 Ethics 567 at 572.

[89] For an example of opportunistic behaviour in long-term supply contracts see *Fratelli Gardino S.p.A.* v. *Caribbean Lumber Co.*, 587 F.2d 204 (5th Cir. 1979). (The seller claims that it is unable to meet contract commitments due to insufficient shipping facilities. The Court finds that facilities were available, but that the seller had chosen to use them for more profitable engagements.)

fairness norms which track those generated by the *fairness* studies discussed earlier.

One model of contract law, employed by Posner, ignores the motives or gains of a person who breaks a contract. Whether the contract breaker fails to perform in order to avoid losses or to generate windfall gains is irrelevant.[90] An alternate model of contract which reflects the gain-loss valuation disparity holds that motives for breach are important, and that failing to perform in order to avoid unanticipated expenses will be treated differently from failing to perform in order to generate a windfall gain to the breaching party.[91] Kessler, Gilmore and Kronman describe the "implicitly amoral"[92] contract theory of Posner and Holmes which ignores the motives for breach, and argue that breaches which are motivated by a gain to the contract breacher will be treated more severely:

> Many will say that [the] breach ought still to be condemned from an ethical point of view ... The inclination to blame A will be even stronger if he has benefitted from his breach, although B has not been harmed ... The Holmesian view ... seems never to have fully overcome the resistance of common sense, which stubbornly insists that moral blame (unlike legal liability) is not entirely a function of the consequences of an action but depends, as well, upon the motives and intentions of the actor.[93]

This second model appears to operate more consistently in contract damage disputes.[94] Predictions,[95] both as to liability and the extent of recovery by plaintiffs, vary directly with information on the motives for breach. In addition, the rule that "motives don't count" has been subject to a number of decisions in which judges have explicitly recognized that they are treating the two situations differently and have designed damage awards deliberately to deprive the breaching parties of the gains which motivated their non-performance.[96]

---

[90] See *supra* note 53; *supra* note 38 at 458-62. See also, *Butler* v. *Fairclough* (1917), 23 C.L.R. 78 at 79; *Asamera Oil Corp.* v. *Sea Oil and General Corp.* (1978), [1979] 1 S.C.R. 633, 89 D.L.R. (3d) 1, (*sub nom. Baud Corp. N.V.* v. *Brook no. 2)* [1979] 3 W.W.R. 93.

[91] See A.S. Burrows, *Remedies for Torts and Breach of Contract* (London: Butterworths, 1987) at 252, 273.

[92] F. Kessler, G. Gilmore & A.T. Kronman, *Contracts, Cases and Materials*, 3d ed. (Boston: Little, Brown, 1986) at 1067.

[93] *Ibid.*

[94] It may operate in many other contexts as well — for example, deliberate breach for profit may influence judicial discretion in saying that a breach is *fundamental* thus triggering judicial review of otherwise enforceable exclusion clauses. Or wilful breach for profit might influence judicial interpretation of terms as conditions rather than warranties. Both choices would work to the contract breachers disadvantage. Perhaps the application of the mistake doctrines can be explained by examining the motives of the person who is seeking to enforce agreements against others who misinterpreted the wodd or the agreement.

[95] What is surprising is that many if not all academics interpret the cases as supporting this view, without any data except their intuitions. Thus Kronman, Gilmore and Kessler suggest that in the case of wilful breach *the inclination to blame [the contract breacher] will be even stronger if he has*

But the treatment of opportunism is more subtle than merely awarding damages to deprive contract breakers of windfall gains. It extends to the interpretation of facts and the discretionary application of principles which permit judges to incorporate *fairness* attitudes in their decision making. These fairness norms, like those reflected in the empirical studies described earlier, are associated with the distinction between gains and losses.

The operation of fairness norms in contract law can be demonstrated by contrasting cases in which judges respond differently depending on whether the otherwise similar behaviour of contracting parties is motivated by attempts to take advantage of unexpected opportunities, or by attempts to respond to unanticipated losses.

One example of such a response might be found in the application of promissory estoppel doctrines in contract performance and modification. Promissory estoppel was developed by the courts in an effort to protect the interests of contracting parties who rely on the unbargained for promises or representations of their contracting parties. It is applied, as most writers admit, to avoid unjust enrichment and the

---

benefited from the breach. See *supra* note 92 at 1067.

[96] For example, in *Penarth Dock Engineering Co. Ltd.* v. *Pounds*, [1963] 1 Lloyd's Rep. 359 (Q.B.) a defendant refused to remove its ship from the plaintiff's dock. Even though the plaintiff did not suffer any loss, the court ordered the defendant to pay money to the plaintiff representing the cost of renting the space which it saved by not having to rent alternate space.

The cases in which judges have explicitly deprived contract breakers of the gains generated by breach are characterized, as the valuation disparity data would predict, by behaviour which often represents deliberate interference with possessory or property rights of the plaintiff. Thus, in *Wrotham Park Estate Co. Ltd.* v. *Parkside Homes Ltd.*, [1974] 1 W.L.R. 798, [1974] 2 All E.R. 321 (Ch.), the defendants entered into a restrictive covenant with the plaintiff which prohibited the construction of houses on the plaintiff's property. The defendant built several houses which did not diminish the value of the plaintiff's property, but were ordered to pay 5% of the 50,000 pounds profit which they made on the construction of the houses. The damages were ordered in lieu of an injunction to demolish the houses. The Court referred to several *tort* decisions which provided for restitutionary awards in the case of interference with the plaintiff's property, and the case has been analogized to that situation. See *supra* note 91 at 274-75.

opportunistic manipulation of legal rules and ambiguous contract language.[97]

A recent example of the application of the doctrine which offers insight into these underlying *fairness* norms is a recent decision of the British Columbia Court of Appeal in a dispute between a landlord and a tenant regarding the ability of the tenant to exercise an option to renew its tenancy.[98] The facts were straightforward. Dukes Cookies had entered into a three-year contractual tenancy agreement with the Alma Mater Society of the University of British Columbia. The lease included a term which permitted Dukes to extend the lease for two years if it gave written notice to the landlord during a specific two-month period.

During the term of the lease, Dukes and the Alma Mater Society began negotiations for a new five-year lease. The negotiations included discussions relating to profit-sharing arrangements and an expansion of the leased premises. The negotiations were prolonged until after the two-month notice period had ended. Dukes failed to give written notice to extend the existing lease because they assumed, based on statements made by the Alma Mater Society's principals, that they would obtain a new five-year lease. After the expiry of the two-month notice period, the Alma Mater Society notified Dukes that they would terminate the lease according to its terms because Dukes had not given written notice to extend the original lease. Dukes brought the legal action to prevent the Alma Mater Society from enforcing the written notice provision in the current lease.

Both the trial judge and the Court of Appeal decided the case in favour of Dukes with the result that it was permitted to exercise its option notwithstanding the expiration of the notice period. The formal justification for the decision was the application of the doctrine of promissory estoppel through which judges can protect the unbargained for reliance of a party to a contract where it would be inequitable to permit the other party to retract a representation which induced the reliance.[99] However, what is obvious from the decision is that the Court could just as easily have decided the case in favour of the landlord on either of two grounds. First, the Court could have found that no representation was made by the landlord to the tenant regarding the likelihood of entering into the new arrangement, as the testimony of both parties was in conflict on this ground. Also, there was considerable authority that there must be an unequivocal and explicit representation in order to trigger the doctrine. Second, the Court could have found that it was unreasonable for the tenant to infer that the notice provision would not be enforced even if the representation had been made.

---

[97] See J.A. Manwaring, "Promissory Estoppel in the Supreme Court of Canada" (1987) 10 Dalhousie L.J. 43 at 51. The common law doctrine of promissory estoppel is reflected in *supra* note 86. See also J.M. Feinman, "Promissory Estoppel and Judicial Method" (1984) 97 Harv. L. Rev. 678.

[98] *Re 6781427 Holdings Ltd. and Alma Mater Society of University of British Columbia* (1987), 44 D.L.R. (4th) 257 (B.C.C.A.), aff'g (1987) 36 D.L.R. (4th) 753.

What is missing from *both* judgments, and yet what must have been critical from the perspective of everyone concerned, is any evidence explaining the abrupt reversal by the landlord of its negotiation position. Why, one can ask, did the landlord negotiate for months as if it would enter into an extended new lease, and then, without warning, give the tenant notice to vacate the premises?

That is the first question which comes to the mind of many readers, and yet is left unanswered in the reported reasons for the decisions. Discussions with the lawyer for the defendant landlord, however, reveal a very different picture. The lawyer indicates that the trial judge was extremely interested in the motives for the landlord's behaviour — an irrelevant consideration in formal terms.[100]. The motive as described in the evidence was simple. The landlord discovered during negotiations that the profits generated by the tenant's business far exceeded its expectations, *and had terminated negotiations in order to open its own cookie business in the same location*! What is remarkable about the case is that neither the trial nor appellate decisions mention this fact. That is, the motive for the termination of the negotiations consisted of an attempt to obtain unexpected gains at the expense of the tenant.

The case can be interpreted as an example of judges responding to the valuation disparity. That is, the unilateral decision by the landlord to terminate the lease without giving the tenant the opportunity to renew for two years generated an expected gain to the landlord *at the expense of the tenant*. The decision can be seen as a reflection of judicial intuition, consistent with the empirical data on the valuation disparity, that the benefit to the landlord, consisting of expected profit, would be substantially less than the harm imposed on the tenant by the eviction. This valuation disparity gives rise to strong perceptions of unfairness consistent with the empirical evidence indicating that when one person gains at the expense of another — a zero-sum game — the transaction is overwhelmingly seen as unfair.[101]

---

[99] The American position is not substantially different, although there is considerable debate regarding the extent to which the promisor can be taken to be responsible for the promisee's reliance. See *supra* note 86, and Feinman, *supra* note 97.

[100] That is, the formal doctrine of promissory estoppel which was applied to justify deciding the case in favour of the tenant, does not include an assessment of the motive of the party who attempts to retract a representation.

A second example of judges responding to the valuation disparity and motive for the defendant's action is a decision on the doctrine of commercial impracticability in *Aluminum Comnany of America* v. *Essex Group, Inc.*[102] That case involved a claim for relief from contractual obligations on the doctrinal grounds of mistake, frustration and commercial impracticability, by the plaintiff Alcoa. Alcoa successfully obtained judicially ordered relief from a long-term obligation to process aluminum ore, which had turned out to be unprofitable because of unanticipated increases in Alcoa's energy costs. The case is generally considered remarkable for two reasons. First, the trial judge was forced to distinguish a long line of cases which most people interpreted as precluding relief where the plaintiff's expected profits had failed to materialize due to unanticipated cost increases. Second, the trial judge did not simply relieve Alcoa from its smelting obligations, but imposed a loss and risk sharing modification to the contract by developing new pricing terms to reduce the losses to Alcoa associated with the explicit contract pricing arrangement.[103]

In *Alcoa* the Court explicitly recognized *fairness* norms as justifying the judicial choices.[104] The problem of course is identifying or articulating the content of this fairness norm. Again, as in *Dukes* the interpretation of judicial behaviour is problematic. But the judgment includes a reference to the formally *irrelevent* fact that when Essex Group had discovered that it was in a position to own substantial amounts of very cheap aluminum, it decided to enter the commodities market. Instead of ordering and using the aluminum in its commercial operations, it had sought out and entered into contracts with third parties to sell its aluminum at current market prices and thus reap a windfall gain at the expense of Alcoa.[105]

---

[101] See *supra* note 37.

[102] 499 F. Supp. 53 (W.D. Pa. 1980) [hereinafter *Alcoa*].

[103] The literature on the case is voluminous, although most writers focus on issues other than the gain-loss distinction which we consider in this paper. See V.P. Goldberg, "Price Adjustment in Long-Term Contracts" (1985) Wis. L. Rev. 527; R.E. Speidel, "The New Spirit of Contract" (1982) 2 *J.L.* & Com. 193; R.E. Speidel, "Court Imposed Price Adjustments under Long-Term Supply Contracts" (1981) 76 Nw. U.L. Rev. 369.

[104] See *supra* note 102 at 76, where the trial judge alludes to the idea that the contract would not be enforced where it would be "commercially senseless and unjust." See also Speidel, "The New Spirit of Contract," *ibid.* at 201 who refers to the courts policing transactions "in the interests

Again, the implication of the judgment is consistent with judicial sensitivity to the valuation disparity which generates perceptions of unfairness. Essex Group was gaining a windfall profit at the direct expense of Alcoa, a transaction which is consistently seen as unfair.

One cannot, of course, point to two ambiguous cases and conclude that judges are responding to conflict only in terms of fairness and the valuation disparity. Certainly there is little explicit recognition of their doing so. What is required to verify the connection of *law* to the fairness data, and what is obviously impossible, is to identify two populations of cases in which all *facts* are identical but for the relevant variable — the obtaining of an advantage at the direct expense of another party. If judges are acting in this way, the outcomes in the two groups of cases should be different.[106]

F. *Repossession*

---

of fairness"; and S.W. Halpern, "Application of the Doctrine of Commercial Impracticability: Searching for the 'Wisdom of Solomon' " (1987) 135 U. Pa. L. Rev. 1123, who refers to an obligation to act in good faith and to agree to contractual adjustments in order to avoid any fortuitous *advantage* at the expense of the other party.

[105] Apparently, Essex Group was taking advantage of its 36.25 cents per pound price for aluminum by selling 25% of its supply in the open market at the then current market price of 73.13 cents per pound. In fact, Essex Group was underbidding Alcoa and according to Alcoa, taking Alcoa's customers. See Note, "Court-Imposed Modifications: Supplementing the All-or-Nothing Approach to Discharge Cases" (1983) 44 Ohio St. L.J. 1079 at 1089 n. 107, 108.

[106] The obvious and insurmountable problem with conducting the experiment is that the two population of cases do not exist. However, we are not alone in suspecting that deliberate breach for profit is influencing the exercise of judicial discretion. For example, John Swan and Barry Reiter, in a discussion of the dissent of Irwin J. in *Peevyhouse* v. *Garland Coal and Mining Co.*, 382 P.2d 109 (Okla. 1963), suggest that he is responding to deliberate decisions to break contracts and thus gain wealth at the expense of another. See *supra* note 80 at 10.

A final example of judges developing principles consistent with the valuation disparity consists of a series of decisions in which judges have restricted creditors' rights to repossess delivered goods on default by buyers prior to full payment of the purchase price. This is evident first, in the presumptive treatment by judges of time of payment clauses — where judges have consistently held that "time is not of the essence" in regard to payment for goods.[107] Absent an explicit contract term, the result is that sellers cannot terminate the contract and repossess goods simply on default, and that buyers are given the opportunity to remedy the defect and retain possession.

Second, even where time of payment clauses have been expressly stated as conditions — thus apparently permitting the seller an election to be discharged from the contract and to repossess the goods — the courts have developed several techniques to protect the possessory welfare of the buyer. These include liberal doctrines of waiver[108] and restrictions on self-help repossession tactics of sellers.[109]

Third, a significant number of jurisdictions have enacted seize or sue legislation which precludes sellers from both repossessing goods and suing for the purchase price.[110] Finally, in many jurisdictions, repossession is permitted only with permission of the court, where the buyer has paid more than two-thirds of the purchase price.[111] This array of judicial and legislative activity can be explained by the data on the valuation disparity which would predict substantial welfare losses associated with repossession of goods from debtors on default.

---

[107] See M.G. Bridge, *Sale of Goods* (Toronto: Butterworths, 1988) at 409. See also, *Mersey Steel and Iron Co.* v. *Naylor Benzon and Co.* (1884), 9 App. Cas. 434, [1881-85] All E.R. 365; *Decro-Wall International S.A.* v. *Practitioners in Marketing Ltd.*, [1971] 1 W.L.R. 361, [1971] 2 All E.R. 216 (C.A.).

[108] See *Pantoutsos* v. *Raymond Hadley Corp. of New York*, [1917] 2 K.B. 473, [1916-17] All E.R. Rep. 448 (the failure to object to late payment held to constitute waiver of the right to treat timely payment as a condition).

[109] For example, sellers must give buyers notice of their intention to repossess or take control of property. See *Traders Bank of Canada* v. *G. & J. Brown Manufacturing Co.* (1889), 18 O.R. 430 (Ch.); *Royal Bank of Canada* v. *Cal Glass Ltd.* (1980), 22 B.C.L.R. 328 (C.A). The restrictions are described in detail in D. Paciocco, "Personal Property Security Act Repossession: The Risk and the Remedy" in M.A. Springman & E. Gertner, eds., *Debtor-Creditor Law: Practice and Doctrine* (Toronto: Butterworths, 1985) at 365. Recent statutory developments have removed the notice requirement. See, for example, *The Personal Property Security Act*, C.C.S.M. c. P35, s. 57.

[110] See *Uniform Commercial Code* § 5.103; *The Limitation of Civil Rights Act*, R.S.S. 1965, c. 103, s. 18; *Sale of Goods on Condition Act*, R.S.B.C. 1979, c. 373, s. 20; *Chattel Mortgage Act*, R.S.B.C. 1979, c. 48, s. 23.

[111] See, for example, *Consumer Protection Act*, R.S.O. 1980, c. 87, s. 23(1).

Schwartz has attacked seize or sue legislation as being "without coherent justification."[112] He argues first, that there is no evidence of systematic underselling of repossessed goods through organized cartels; and second, that repossession does *not* impose greater harms on debtors than it creates gains for creditors. Schwartz does discuss the valuation disparity, but unfortunately he misses its significance in several ways. First, in discussing the effects of repossession, he admits that the losses associated with repossession may reflect the fact that debtors may value the goods more than the market price. However, he then states that the "harms could not occur if debtors were perfectly informed of the consequences of granting security."[113] That is, if the future welfare loss were known at the time of contracting, then the debtor would either not grant security, or would grant less, and pay a higher interest rate. Whether Schwartz is correct or not depends on when the buyer's reference point changes, and thus when the repossession is perceived as a loss rather than a foregone gain. If, as seems reasonable, the shift does not take place until some time after possession of the good, then the buyer's ability to contract, *ex ante*, is compromised.

Second, even assuming debtor ignorance of the costs of repossession, Schwartz argues that the alleged welfare losses associated with repossession either would not occur or would be trivial. He acknowledges that if debtors value the goods because they own them, and thus if the amount they would demand to give up their goods exceeds, in a significant way, the amount which they would pay to obtain the goods, then the assignment of legal rights could determine outcomes more frequently than is commonly supposed. However, the valuation disparity may be linked to a kind of relationship connected not to the assignment of legal rights, but to possession. If that is so, then merely shifting *ownership* will not eliminate the valuation disparity.

Third, in assessing whether repossession generates losses, he denies the empirical validity of claims that "creditors necessarily derive less value from repossessed goods than debtors lose."[114] While the empirical studies described above do not support the proposition that the creditors *necessarily* derive less value, they certainly do offer support for the proposition that the loss to the debtor will consistently, and in a very significant way, exceed the expected gain to the creditor.

---

[112] A. Schwartz, "The Enforceability of Security Interests in Consumer Goods" (1983) 26 J.L. Econ. 117 at 161.

[113] *Ibid.* at 140.

[114] *Ibid.* at 142-43.

Finally, Schwartz argues that the gain of the buyer who purchases repossessed goods from the creditor may offset the loss to the original buyer, and thus that repossession does not necessarily destroy value. Again, the empirical studies discussed earlier suggest that the loss of economic welfare measured by the original buyer's reservation price will almost certainly exceed the gain of the second buyer.

Our conclusion, drawn from the empirical evidence which we describe above, is that buyers in possession will consistently value the possessed goods significantly above the value placed on them either by repossessing creditors or hypothetical future buyers, and the purchaser will not generally be able to take the higher valuation into account when contracting. Recognizing this valuation disparity offers considerable support for the judicial restrictions on repossession framed in common law rules and presumptive interpretations of contract language, as well as for the legislative restrictions on repossession adopted in many jurisdictions.

IV. CONCLUSION

Contrary to the assertions of conventional economic practice and the prescriptions of many critics of legal outcomes, people commonly do not value losses the same as they do gains. The greater weight given to losses over objectively commensurate gains appears pervasive and is a major determinant of what actions or changes meet community standards of being fair or acceptable and which are likely to be considered unfair or less acceptable.

Further, legal institutions appear to reflect the same disparity between valuations. Court decisions, over a very broad array of cases, seem to take account of these differences in according greater protection of losses over foregone gains.

This greater weight given to losses in legal outcomes is not only consistent with people's judgments of fairness and normal business practices, but to the extent that these different weights reflect actual valuations of welfare changes, this differential treatment in judicial choices may also promote efficiency. Equivalent treatment of gains and losses called for by reform proposals "to reconcile" the differences[115] seems not only unlikely to be adopted, but may also lead to inefficient outcomes and be contrary to community fairness and equity standards.

---

[115] Waddams, *supra* note 7 at 540.

ATTACHMENT G

VALUATION OF GAINS AND LOSSES, FAIRNESS AND NEGOTIATION OUTCOMES

Bernhard F. J. Borges and Jack L. Knetsch
Simon Fraser University

Mail should be addressed to:

Jack L. Knetsch
School of Resource and Environmental Management
Simon Fraser University
Burnaby, British Columbia
Canada V5A 1S6

VALUATION OF GAINS AND LOSSES, FAIRNESS AND NEGOTIATION OUTCOMES

Bernhard F. J. Borges and Jack L. Knetsch
Simon Fraser University

Abstract

This paper reports the results of two experimental tests of the extent to which the large disparity between people's valuation of gains and losses, and related fairness determinations, are used in judging the acceptability of alternative negotiating or conflict resolution proposals. Participants acted as arbitrators and selected their preferred resolution of conflicts involving either the division of gains or sharing responsibility for losses. Different cases were presented in which one or the other party incurred varied combinations of direct or opportunity costs, or received varied forms of payments.

The patterns of resolutions were fully consistent with earlier behavioural findings. Contrary to conventional economic assumptions, direct costs incurred by one party to the negotiation were far more important than opportunity costs in setting the terms of a more acceptable resolution.

The results strongly suggest that recent behavioural findings might be used to improve the design of negotiating and conflict resolution proposals.

VALUATION OF GAINS AND LOSSES, FAIRNESS AND NEGOTIATION OUTCOMES[1]

Bernhard F. J. Borges and Jack L. Knetsch
Simon Fraser University

Conventional economic assumptions of motivation, preferences and behaviour are often used in the design of negotiation and dispute resolution proposals. Numerous recent studies have found, however, that many behavioural assertions used in traditional economic analyses frequently provide neither very good descriptions of people's preferences nor very useful predictions of what economic activities they find acceptable. For example, empirical evidence suggests people are influenced by sunk costs, value losses far more than commensurate gains, take much greater account of out-of-pocket than opportunity costs, and commonly judge the acceptability of economic actions by the perceived fairness of the resulting consequences.

Just as many public policies and private actions could benefit by taking greater account of these findings of people's preferred choices, proposals to resolve conflicts and obtain negotiated agreements might be improved and gain greater acceptance by exploiting this evidence. This paper reports the results of two tests of the extent to which the behaviour of people observed in experimental and market settings, as opposed to traditionally assumed behaviour, might influence their judgments of the acceptability of alternative conflict resolution proposals.

---

[1] This research was in part supported by the Social Science and Humanities Research Council of Canada, the Science Council of British Columbia and the Ontario Ministry of the Environment and has benefited from the suggestions of Shane Frederick, Andrew Gemino, Robin Gregory and Mark Wexler.

Specifically, the study tested the extent to which people may be more sensitive to the widely observed disparity between people's valuation of gains and losses than they are to conventional assumptions of gain and loss equivalence, in their judgments of the acceptability of alternative resolution offers.

Two experiments were conducted in which participants acted as arbitrators charged with selecting a preferred resolution of a series of disputes in which two negotiating parties had incurred varying combinations of direct and opportunity costs. The first experiment involved parties having conflicting interests in the sharing of a gain; the second experiment concerned disputes over sharing responsibility for a loss. The findings suggest that weighing losses more than gains and taking more account of direct than opportunity costs, which have been observed in other contexts, are also strongly reflected in what individuals regard as fair and acceptable resolutions of conflicts. These results lend strong support to expanding the widely advocated negotiation strategy of appealing to community standards or objective criteria (e.g. Fisher and Ury, 1981), to include such persistent preference patterns and to rely less on assumed behaviour that is inconsistent with these preferences.

Normative Prescriptions and Observed Behaviour

Among the standard economic principles used in attempts to resolve conflicts are those prescribing that bygones be ignored, and opportunity and direct costs be treated equivalently: "... sunk costs should have no influence on deciding what is currently the most profitable thing to do" (Lipsey, Parvis and Steiner, 1985, p. 149); " ... a receipt foregone of a given amount is the equivalent of a payment of the same amount" (Coase, 1960, p. 7). Further, any questions of equitable or acceptable distribution of gains and losses are largely left to strictures that factors be awarded the value of their marginal productivity -- which in terms of equity theory reflects input-to-output ratios of the parties (Adams, 1963, 1965; Walster, Walster and Berscheid, 1978) -- and vertical equity judgments that the rich ought not to benefit at the expense of the poor, based largely on the notion that the marginal value of a gain or loss likely larger for a poor person. Beyond this there is usually a strong preference for equal division of gains or losses, such as those that might be at issue in a negotiation (Frank, 1988) -- though defining the appropriate subtractions to derive the net change is often contentious (e.g., Reekie, Allen and Crook, 1991).

While conventional views of economic principles may remain the basis of much analysis and prescription, recent empirical evidence clearly points to pervasive systematic departures of actual from predicted choices. In part, these observed differences stem from a series of cognitive limitations that lead to employment of rules of thumb, or heuristics, that result in consistently biased choices (Tversky and Kahneman, 1973; Kahneman, Slovic and Tversky, 1982; Bazerman, 1986; Neale and Bazerman, 1991; Bazerman and Neale, 1992,).

Beyond such cognitive biases, however, are further instances in which basic assumptions of commonly applied economic models are seriously at variance with the observed preferences and behaviour of people. An important example is the disparity between people's valuation of losses relative to fully commensurate gains (Kahneman and Tversky, 1979; Kahneman, Knetsch and Thaler, 1991). Contrary to the usual assumption of equivalence, long a staple of conventional economic practice and analysis, numerous studies of people's preferences and market and non-market choices find that losses from a reference position are valued far more than gains to it (Kahneman, Knetsch, and Thaler, 1990).

The earliest reports of valuation disparities were based on the results of survey respondent's assessments of various environmental changes and first appeared some twenty years ago (Hammack and Brown, 1974). Respondents invariably indicated that they would demand a much larger sum to agree to the demise of an amenity than they would be willing to pay to preserve it. More compelling real exchange studies were initiated a decade later and revealed similar differences (Knetsch and Sinden, 1984). The sums that subjects agreed to pay for a variety of assets, such as lottery tickets and a range of small consumer goods, were much less than the amounts of money they would demand to give up these same entitlements (Knetsch, 1989). Later tests using a variety of study designs have provided evidence that the difference between gains and losses is not eliminated by repeated valuations, is not due to strategic behaviour and is independent of wealth effects and transaction costs (Kahneman et al., 1990). Reports of valuation disparities by many investigators, including some who earlier reported contrary views, are now becoming quite commonplace in leading economics journals. While many of the reported tests are based on comparisons of the buying or selling decisions of different groups of subjects, similar disparities have also been shown in the buying and selling decisions of the same individuals (Kachelmeier and Shehata, 1992; Dubourg, Jones-Lee and Loomes, 1993).

Valuing losses more than otherwise fully commensurate gains appears, on the basis of the available empirical evidence, to be a pervasive characteristic of people's preferences. Further, the large difference in the valuations of gains and losses implies that opportunity costs, viewed as foregone gains, will be valued less than out-of-pocket costs, which are more likely to be treated as losses (Thaler, 1980).

Judgments of fairness and acceptability are also more consistent with the gain/loss disparity findings than they are with traditional assertions of economic motivations. Survey and experimental studies have found that actions imposing losses on particular parties are widely regarded as more unfair than ones which result in foregoing gains (Kahneman, Knetsch and Thaler, 1986a and 1986b). For example, cutting wages of workers in the absence of a firm losing money was widely judged to be unfair, but reducing employees' yearly bonuses by the same amount was seen to be fair and acceptable. A wage cut is apparently viewed as an imposition of a loss, whereas a bonus reduction is framed as a less serious foregone gain. For similar reasons, people overwhelmingly indicated that raising a price above the current list price is more unacceptable than a comparable increase brought about by terminating a current discount (Kahneman, Knetsch and Thaler, 1986a).

These results were generally replicated by Gorman and Kehr (1992) using a U.S. managerial sample; by Shiller, Boycko and Korobov (1991), in a comparison of New York and Moscow residents, and by Frey and Pommerehne using a Swiss and German household sample (1993).

Guth, Schmittberger and Schwarze (1982) showed that procedural justice played a significant role in the outcome determination of a series of two-player ultimatum games. Parties often rejected offers unless a fair, or equal, division of the stakes was proposed -- individuals effectively gave up unacceptable, yet positive payoffs, in order to punish their opportunistic opponents. Overall, the empirical results display a systematic pattern suggesting a positive covariance between decision makers' loss aversion and their subsequent judgment of fairness. This result appears to reflect widespread community standards, or social norms, that largely govern the judgment of what is and is not acceptable, or fair, behaviour.

Further, judicial choices in common law jurisdictions appear to reflect much more sensitivity to the gain/loss disparity than most conventional economic analyses would prescribe. Recovery of expenses resulting from wrongful actions is far more likely than compensation for foregone profits, for example. Consistent with a heavier weighing of losses over gains, courts also show a clear pattern of reluctance to move entitlements from current holders to another party (Cohen and Knetsch, in press).

Tests Of Direct And Opportunity Costs On Arbitrators' Choices

Evidence from valuation and fairness studies indicates not only that individuals tend to take full account of direct costs and to neglect opportunity costs in everyday economic decisions, but also that they consider it 'fair', or acceptable, for others to do so. In two experiments, conducted to test the applicability of this asymmetric treatment of direct and opportunity costs in judging the acceptability of negotiation and conflict resolution proposals, subjects (Ss) acted as arbitrators and were asked to select what they believed to be the most acceptable resolution of each of a series of disputes between two parties. To better assure that choices made by subjects reflected general, or community, standards of acceptance rather than influences of personal gain, participation in each experiment was voluntary and subjects were offered no compensation or other benefit that might influence their decisions.

Experiment 1 - Costs and Division of Gains

Fourtyone (41) fourth year business students enroled in strategic management seminars at Simon Fraser University were approached during class time and requested to participate in the experiment. The average age of the participants was 24 years: 42 per cent were female and 68 per cent cited English as their first language.

Each participant was asked to assume the position of an arbitrator required to settle five disputes involving the division of a $10 gain between two people. Ss were given a multiple choice questionnaire (see Appendix 1) and requested to simply select, from the options made available to them, the allocation of the $10 between the two negotiators from the available option which they considered most suitable for each of five cases.

Case 1: The arbitrators were told that the (two) parties to the negotiation were to split a $10 gain between the two of them, and that neither party had incurred any cost associated with the exercise. Each subject was asked to select one of two options as their recommended resolution: (a) a $5/$5 division of the $10 and (b) a noncooperative $10/$0 'division'. This case provided a baseline for Ss' fairness-preference in this context.

Case 2: In this case, Ss were told that one party was required to pay an out-of-pocket cost of $2 in order to take part in the division of the $10, whereas the other party did not incur such a cost. Three choices for resolving the dispute were offered: (a) a $5/$5 split, (b) a $6/$4 choice that took account of the direct cost to the one party, and (c) a noncooperative $10/$0 choice. The $6/$4 offer was supplemented with a cue concerning the arithmetic associated with the choice, i.e., "To account for one party's $2 cost."

Case 3: Rather than a direct cost, as in the second case, this scenario stated that one party faced a $2 opportunity cost, or foregone gain, in the form of the individual having given up an alternative exercise which would have resulted in receiving $2. The choices provided to the arbitrators were identical to the ones' provided for the second case, namely, a $5/$5, a $6/$4 and a $10/$0 division. Only the $6/$4 choice was 'cued', i.e., "To account for the other party's foregone gain."

Case 4: Here, one negotiator was said to have incurred a $2 out-of-pocket cost, while the other negotiator benefited from a $6 windfall profit. The point of this scenario was to examine the extent to which a windfall profit accrued by one of the parties is considered to be extraneous and should or should not be taken into account on grounds that it is or is not irrelevant to the task at hand. A total of five choices were offered to Ss: a $5/$5 offer, a $6/$4 offer that took account of the cost but ignored the windfall, a $8/$2 offer that took note of only the windfall profit, an offer of $9/$1 reflecting equal sensitivity to the cost and profit, and a noncooperative $10/$0 offer. With the exception of the $5/$5 and the $10/$0 offer, all choices were supplemented with explanatory cues similar to the wording employed for Case 2 and Case 3.

Case 5: For this last arbitration task, one negotiator was said to have incurred a $4 opportunity cost, whereas the other party was required to pay a $2 direct cost. Arbitrators were offered five choices: a cost-neutral $5/$5, a $3/$7 choice that ignored the direct cost but was sensitive to the opportunity cost, a direct and opportunity cost integrative $4/$6, a direct out-of-pocket integrative cost but opportunity cost neutral $6/$4, and a $10/$0 offer. Again, cues were provided for all choices but the $5/$5 and the $10/$0 offers.

Results

Neither gender nor English as a first language, revealed any significant effect on arbitrators' choices. Also, Ss were asked what they thought the experiment was all about and whether this guess had influenced their responses. This manipulation check did not reveal any significant effect for any of the five responses.

Nine of the 41 (22%) Ss invariantly opted for $5/$5 splits in all five cases. The motivations likely varied, possibly including this being a means of getting through quickly. In a debriefing session following the experiment, however, most of these participants indicated that they ignored any kind of costs or profits as they considered these to be sunk, " ... and thus irrelevant". The results with these observations included and excluded are both of interest -- to some extent they may reflect specific criteria of fairness and acceptability, and to some extent they may unduly inflate the cost-neutral response. The frequency distributions for the five negotiation tasks are summarized in TABLE 1, with the percentage of Ss excluding the invariant $5/$5 responses given in parentheses below the percentage derived from the total sample.

The baseline first case called for the division of $10 with no costs incurred by either negotiator. An even split was clearly regarded as 'the right thing to do', as all Ss selected this rather than the opportunistic $10/$0 choice.

The second case revealed that arbitrators thought that the $2 out-of-pocket, or direct, cost should be taken into account in determining an acceptable division of the $10. Even with the inclusion of the invariant choosers of the $5/$5 split, three out of four (76%) recognized the $2 as a cost to be included in the determination of negotiators' fair net returns; and with the constant $5/$5 choosers excluded, nearly all of the remaining arbitrators (97%) settled for the $6/$4 offer.

In contrast to the accounting for direct cost, the third case showed that opportunity costs were not considered as appropriate determinants of a fair division of the $10. More than four out of five (83%) of the arbitrators opted for the opportunity cost neutral $5/$5 division (and 78%, excluding the invariant $5/$5 choosers) suggesting that a party's foregone gain should not enter into the determination of appropriate rewards.

In the fourth scenario, where one negotiating party received a $6 windfall and the other incurred a $2 direct cost, the majority of the arbitrators (61%) opted for the $6/$4 division (and 78%, excluding the $5/$5 choosers) acknowledging the direct cost but ignoring the windfall profit. Twentynine per cent (9% without the $5/$5 people) chose to ignore both the direct cost and the windfall profit, whereas 10% (or 13% without the $5/$5s) of the respondents opted to account for both.

The fifth and last scenario tested for the impact of both types of costs on fairness standards, as arbitrators were told that one negotiator had incurred a $2 direct cost and the other negotiator a $4 opportunity cost. Again, the majority of the arbitrators (56%; but 72% without the $5/$5s) acknowledged the direct cost while ignoring the foregone gain in their decision. Another 29 percent (or 9% without the $5/$5s) of the Ss chose to ignore both the direct cost and the foregone gain by choosing a $5/$5 division; whereas 15% (or 19% without the $5/$5s) of the Ss decided to account for both cost types and deemed the $4/$6 settlement appropriate.

*** Insert TABLE 1 Approximately Here ***

Experiment 2 - Costs and Division of Losses

One hundred and thirtyfour (134) second and third year business students enroled in second and third year management courses at Simon Fraser University participated in the second experiment. The average age of the participants was 22 years; 63 per cent were female and 43 per cent claimed English as their first language.

Each participant acted as an arbitrator required to settle a dispute involving two people dividing responsibility for a $10 *loss* between them. The arbitrators were again asked to select an allocation of the $10 loss which they considered most suitable. A multiple choice questionnaire comprised of four questions was employed and monetary divisions that were potentially arithmetically demanding were supplemented with a rationale, i.e., 'cued'.

Although the four scenarios were parallel to four of the five cases employed in the first experiment,[2] the cost positions incurred by the negotiators were expanded to further examine the sensitivity to the gain/loss valuation disparity. That is, the second experiment encompassed various combinations of direct and opportunity costs as well as direct gains and avoided losses (i.e. "opportunity gains"). A partial factorial design was employed which required two versions of questions 2 and 3, and four versions of question 4 (see Appendix 2 for an example of the questions used for this experiment).

The first case, which was identical for all four questionnaires, reflected the 'costless' $5/$5 scenario. Specifically, arbitrators were asked how they would allocate a $10 loss between two negotiating parties.

---

[2] For reasons of brevity the exogenous 'windfall' scenario, i.e. scenario four, of the first experiment was not included in this second experiment.

The second case corresponded closely to the second case of the first experiment, but consisted of two variations. The first variation had one negotiating party incur a $2 direct cost, whereas the second version had one negotiating party benefit from a $2 payment. Each variation was employed in two of the four questionnaires.

The third scenario was also split into two variations. Instead of a direct cost and gain, however, one negotiation party either (i) incurred an opportunity cost of $2, which she or he would have earned by participating in another lottery-type exercise, or (ii) avoided losing $2 in the alternative lottery exercise. Again, the negotiating opponents did not incur, or benefit, from either such an opportunity cost or gain.

Finally, the fourth case, corresponding to the fifth case of experiment one, consisted of four versions representing an asymmetric combination across the two negotiators of the four types of cost: (i) one party paid $2, the other incurred a $4 opportunity cost; (ii) one party paid $2, the other avoided a $4 loss; (iii) one party gained $2, the other incurred a $4 opportunity cost; and (iv) one party gained $2, the other avoided a $4 loss. The rationale for the $4/$2 differential sums is the same as in Experiment 1.

Results

English as a first language did not reveal any significant effect on any of the arbitrators' choices, nor did gender with the single exception of the first case ($p < 0.1$, $n = 63, 71$, 1 df, K-W tests statistic = 3.78).

As in the first experiment, and presumably for similar motives, 22 per cent (30) of the Ss chose $5/$5 offers for all four cases. For comparison with and without their inclusion, the frequency distributions for the various negotiation tasks are summarized in TABLE 2a and TABLE 2b respectively.

In the baseline first case which called for the division of the \$10 loss with no costs incurred by either negotiator, nearly all of the arbitrators, 97 per cent (and 96% excluding the invariant -\$5/-\$5 Ss) , opted for an equal, -\$5/-\$5 allocation of the loss between the two parties. This is not significantly different from the results of the baseline case of the first experiment for both the unadjusted and the adjusted data.[3]

The second case, involving direct costs or gains, revealed that arbitrators clearly took these into account. Seventy-three per cent of the arbitrators thought that the out-of-pocket costs incurred by one of the parties should matter to the allocation decision, and 60% of the arbitrators took account of the \$2 gain paid to one of the parties. (Omitting the invariant -\$5/-\$5 responses, 86% of the Ss opted to account for the \$2 cost, whereas 85% considered the \$2 gain relevant to their arbitration decision.) While more of them took note of costs, there was no statistically significant differentiation between the direct cost and receipt of a gain in their allocation decisions. Both direct costs and gains were treated as a 'tangible' cash flow, and therefore considered fair, or acceptable, to take fully into account in settlements. There was no significant diffference between the Case 2 results of the two experiments -- direct costs were taken into nearly equal account for both gains and losses[4].

---

[3] $p > 0.1$, $n = 134, 41$, 1 df, K-W test statistic = 0.31 and $p > 0.10$, $n = 104, 32$, 1 df, K-W tests statistic = 0.31, respectively.

[4] Comparisons between the first and second experiment are based on recoded data to adjust for the value asymmetry associated with the differential gain/loss context of the two experiments.

The third case showed that neither opportunity costs nor avoided losses were considered appropriate determinants of a fair division of the $10 loss. More than four out of five of the arbitrators, (85%), of the arbitrators selected the opportunity cost neutral $5/$5 division. The conceptually identical 'avoided loss' condition, however, was recognized as irrelevant by a smaller proportion (69%).[5] Excluding the invariant -$5/-$5 responses, 78% ignored the foregone gain and 61% disregarded the avoided loss -- a significant difference (K-W test statistic = 3.45, p <0.1,).[6]

There is a large and significant difference between the allocations prescribed in the second and third cases for both the gains and cost conditions for both the unadjusted and the adjusted data sets ($p < 0.005$ for all four comparisons). This strongly supports the conclusion drawn from the first experiment that direct costs, or gains, are valued much more than the conceptually equivalent opportunity costs/avoided gains, and that negotiation proposals are judged accordingly[7].

The fourth and last scenario tested for the impact of all four combinations of direct costs/gains and opportunity costs/avoided losses. In the first, of one negotiator having incurred a \$2 direct cost and the other a \$4 foregone gain. 36 per cent of the Ss ignored both costs, but 42 per cent acknowledged the direct cost only. The response frequencies for the \$4 opportunity cost/\$2 gain condition were not significantly different (z-stat = 0.38): 46 per cent disregarded both and 40 per cent accounted for the \$2 gain but ignored the foregone gain. .

---

[5] The difference is not significant; but it should be noted that p=0.1138 and is only marginally so.

[6] Although both foregone gains and avoided losses were not taken into account by the large majorities of Ss, the smaller proportions doing so for avoided losses in this case, as well in the previous case 2, indicates a greater concern for losses relative to gains that is consistent with both general findings of gain/loss value disparities as well as the results obtained for the first experiment.

[7] Again, there was no significant differences in arbitrators' decisions between the first and second experiment.

Excluding the invariant -\$5/-\$5 respondents shows, of course, an even stronger pattern of worrying about direct costs and gains and neglecting foregone gains. Over half (52%) did so for costs, and slightly more (56%) for gains -- again, an insignificant difference (z-stat = 0.66).

The remaining two endowment conditions pitted negotiators having avoided a \$4 loss against negotiators that either benefitted from a \$2 gain or incurred a \$2 direct cost. For the avoided loss/direct cost condition, 51 per cent of the arbitrators accounted for the gain but ignored the avoided loss. However, while not a significant difference (z-stat = 2.12, $p > 0.1$), only 21 per cent of the arbitrators accounted for the gain only in the avoided loss/gain condition, while 33 per cent ignored both cost types and 33 per cent of the arbitrators accounted for both -- in essence equating the loss, even though only an avoided one, with an actual gain!

Adjusting for the invariant responses further substantiates these observations. Approximately half of the arbitrators (48%) accounted for only the gain in the gain/avoided loss condition, and 58 per cent took account of only the direct cost in the avoided loss/direct cost condition.

There was no significant difference between the last scenario of the first experiment, in which one party incurrs a \$4 opportunity cost and the other a \$2 direct cost, and either the foregone gain/direct cost or foregone gain/gain versions of the last question of the second experiment. There were, however, significant differences for the avoided loss/direct cost ($p<0.05$,

$n = 31, 41$, 1 df, K-W statistic = 54.19) and avoided loss/direct gain ($p < 0.05$, $n= 33, 41$, 1 df, K-W statistic = 9.25) versions. These results further suggest that even though direct gains or actual payments dominate either foregone gains or avoided losses in judging the acceptability of dispute settlements, avoided losses were weighed more heavily by arbitrators than foregone gains.

*** Insert TABLE 2a and 2b Approximately Here ***

Discussion

The pervasive evidence that people value losses more than commensurate gains implies that foregone gains will be commonly weighted less than losses, and consequently, that direct or out-of-pocket costs will be more important than opportunity costs. The present study has shown that people not only distinguish between direct and opportunity costs in their buying and selling decisions and in other economic choices, but also view them differently in judging the acceptability of negotiating positions. Furthermore, the experiment took place in business administration classrooms with participation by lower and upper division business students, thereby deliberately encouraging symmetric treatment of costs in the tests.

Although the evidence from these experiments is limited to two-party negotiations over a simple division of a sum of money, it presents a clear demonstration that 'elusive' costs, such as opportunity costs and avoided losses, have less legitimacy in negotiation positions than direct costs incurred by one of the parties. This result suggests that proposals to resolve conflicts that take this into account would in all likelihood be more acceptable to all parties than ones based on conventional views of gain/loss equivalence.

Just as loss aversion gives rise to the greater importance of losses over foregone gains examined in this study, the gain/loss disparity also suggests that reducing a loss will typically be valued much more than an otherwise commensurate gain. The salience of the reference position will therefore become important in disputes over such things as workplace safety or community security where changes may be viewed either as reductions of hazards or gains in safety, or weighing the gains in air quality which can be taken either as an improvement or reduction of pollution. More generally, concessions that reduce losses will be favoured -- and better promote settlement -- over ones that provide gains (Kahneman and Tversky, in press).

This study demonstrates but one case of taking advantage of the evidence that people's negotiation judgments differ from those prescribed by conventional assertions of economic motivation. It is likely that the same could be done with others. The different weight commonly accorded changes that provide certainty over near certainty, or ones that introduce risk to a previously certain environment, provide other opportunities to build on available evidence of observed rather than assumed preferences of people. While standard economic principles treat similar changes in the probability of an event of equal significant regardless of initial likelihood, people invariably demonstrate a very different weighting of certainty and moves that achieve or deny it -- the so-called "certainty effect" (Kahneman and Tversky, in press). Negotiating positions that include a guarantee of something happening or a guarentee of precluding it, are likely to be taken much more seriously than ones which offer commensurate change but fall short of this threshold (Neale and Bazerman, 1991).

Other areas in which people's judgments of fairness do not readily fall out of traditional views of economic motivations present further potential possibilities for improving the design of proposals to resolve conflicts. For example, contrary to common assertions, distributional worries do not centre on vertical equity; instead, people appear at least as concerned with worries of horizontal equity, or equal treatment. In differences over tenancy arrangements it is widely agreed that it is fair for a landlord to pass on the increased costs associated with the maintenance and operation of a rental property to a tenant, quite independent of the relative wealth positions of the two parties. Similarly, differences in the incomes of the two do not mitigate the widespread view that it is not fair to increase rents to cover other losses incurred by a landlord (Kahneman, Knetsch and Thaler, 1986a and 1986b). Negotiation positions that reflect such distinctions are more likely to gain support than ones that ignore them.

The evidence of fairness criteria might also prove beneficial to understanding the often great reluctance of workers to accept wage reductions. It is widely seen as unfair for profitable firms to impose losses on incumbent workers by cutting their wages in order to gain further profits. However, such unfair moves of one party gaining at the other's expense are generally made acceptable when this zero-sum characteristic is eliminated by new employees being offered a lower wage or by a firm using a wage reduction to reduce its losses.

The reluctance of workers to accept wage roll-backs, for airplane or ferry crews to accept cuts in their numbers, for beligerants to accept mutual reductions in their arms, or for professional bodies to accept a wider sharing of their prerogatives are but instances of a general *concession aversion* (Kahneman, 1992) that become more understandable in the light of evidence of gain/loss valuation differences, influences on fairness judgments, and similar behavioural characteristics. Continued adherence to conventional views of economic self-interest are likely to lead to missed opportunities for improving the design of negotiation and conflict resolution proposals that could be yielded by a more realistic view of what matters to people.

References:

Adams, J. (1963). Toward an understanding of inequity. *Journal of Abnormal and Social Psychology*, 67, 422-436.

Adams, J. (1965). Inequity in social exchange. in: Berkowit.z, L. (ed.) *Advances in Eexperimental Social Psychology*, 2, 267-299. New York: Academic Press.

Bazerman, M. H. (1986). *Judgment in decision making.* New York: John Wiley and Sons.

Bazerman, M. H., & Neale, M. A. (1992). *Negotiating Rationally*. New York: The Free Press.

Coase, R. (1960). The problem of social cost. *Journal of Law and Economics*, 3. 1-44.

Cohen, D., & Knetsch, J.L. (in press). Judicial choice and disparities between measures of economic values. *Osgoode Hall Law Journal.*

Dubourg, W. R., Jones-Lee, M.W., & Loomes, G. (1993). The contingent valuation of transport safety: imprecise preferences and the disparity between willingness to pay and willingness to accept. Centre for Social and Economic Research on the Global Environment (CSERGE) Discussion Paper, UK.

Fisher, R., & Ury, W. (1981). *Getting to Yes: Negotiating Agreement Without Giving In.* New York: Houghton Mifflin.

Frank, R.H. (1988). *Passions within reason.* New York: W.W. Norton and Company.

Frey, B.S., & Pommerehne, W.W. (1993). On the Fairness of Pricing -- an Empirical Survey Among the General Population. *Journal of Economic Behavior and Organization*, 20, 295-307.

Gorman, R.F., & Kehr, J.B. (1992). Fairness as a constraint on profit seeking: comment. *American Economic Review*, 82, 355-358.

Guth, W., Schmittberger, R., & Schwarze, B. (1982). An experimental analysis of ultimatum bargaining. *Journal of Economic Behaviour and Organization*, 3, 367-380.

Hammack, J., & Brown, G. (1974). *Waterfowl and wetlands: toward bioeconomic analysis.* Baltimore: John Hopkins Press.

Kachelmaier, S., & Shehata, M. (1992). Examining risk preferences under high monetary incentives: experimental evidence from the Peoples' Republic Of China. *American Economic Review*, 82, 1120-1141.

Kahneman, D. (1992). Reference points, anchors, norms, and mixed feelings. *Organizational Behavior And Human Decision Processes*, 51, 296-312.

Kahneman, D., Knetsch, J.L., & Thaler, R. (1986a). Fairness and the assumptions of economics. *Journal of Business*, 59, 285-300.

Kahneman, D., Knetsch, J.L., & Thaler, R. (1986b). Fairness as a constraint on profit seeking: Entitlements in the market. *American Economic Review*, 76, 728-741.

Kahneman, D., Knetsch, J.L., & Thaler, R. (1990). Experimental tests of the endowment effect and the Coase theorem. *Journal of Political Economy*, 98, 1325-1348.

Kahneman, D., Knetsch, J.L., & Thaler, R. (1991). The endowment effect loss aversion and status quo bias. *Journal of Economic Perspectives*, 5, 193-206.

Kahneman, D., Slovic, P., & Tversky, A., eds. (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge, Ma.: Cambridge University Press.

Kahneman, D., & Tversky, A. (1979). Prospect Theory: An analysis of decision under risk. *Econometrica*, 47, 263 - 291.

Kahneman, D., & Tversky, A. (in press). Conflict resolution: a cognitive perspective. In K. Arrow et al. (eds.). *Barriers to conflict resolution*. New York: Norton and Co.

Knetsch, J.L. (1989). The endowment effect and evidence of non-reversible indifference curves. *American Economic Review*, 79, 1277 - 1288.

Knetsch, J.L, & Sinden, J.A. (1984). Willingness to pay and compensation demanded: experimental evidence of an unexpected disparity in measures of value. *Quarterly Journal Of Economics*, XCIX, 507 - 521.

Lipsey, R. G., Parvis, D. D., & Steiner, P. O. (1985). *Economics*. Fifth Edition. New York: Harper and Row Publ.

Neale, M.A., & Bazerman, M. H. (1991). *Cognition and rationality in negotiation*. New York: Free Press.

Reekie, W.D., Allen, D.E. and Crook, J.N. (1991). *The economics of modern business*. Second Edition. Cambridge, Ma.: Blackwell Publishers.

Schiller, R. J., Byocko, M.,& Korbov, V. (1991). Popular attitudes toward free markets: the Soviet Union and the United States compared. *American Economic Review*, 81, 385-400.

Thaler, R. H. (1980). Toward a positive theory of consumer choice. *Journal of Economic Behavior and Organization*, 1, 39-80.

Tversky, A., & Kahneman, D. (1973). Availability: a heuristic for judging frequency and probability. *Cognitive Psychology*, 5, 207-232.

Walster, E., Walster, G., & Berscheid, E. (1978). *Equity: theory and research.* Boston: Allyn & Bacon.

TABLE 1: Response Frequencies For Experiment 1 - Division of Gains[a]

	$3/$7	$4/$6	$5/$5	$6/$4	$8/$2	$9/$1	$10/$0
Case 1:			100% **(100%)**				0%
Case 2			24% (3%)	**76%** **(97%)**			0%
Case 3			**83%** **(78%)**	17% (22%)			0%
Case 4			29% (9%)	**61%** **(78%)**	0% 0%	10% (13%)	0%
Case 5	0% 0%	15% (19%)	29% (9%)	**56%** **(72%)**			0%

[a]: Choices in accord with the proposed community standard appear in bold. Frequencies for the 'adjusted' data set are captured by the second cell entry

TABLE 2a: Unadjusted Frequencies - Experiment 2[a] (Losses)

	0	2/8	3/7	4/6	5/5	6/4	7/3	8/2	10/0
Q1 (n = 134)	0%			1%	**97%**	2%			0%
Q2									
Direct Cost (n = 66)	3%			**73%**	23%				2%
Gain (n = 68)	0%				37%	**60%**			3%
Q3									
Foregone Gain (n = 66)	2%			9%	**85%**				6%
Avoided Loss (n = 68)	0%				**69%**	28%			1%
Q4									
Foregone Gain/Direct Cost (n = 31)	0%			**42%**	36%	13%	10%		0%
Avoided Loss/Direct Cost (n = 35)	0%	23%	6%	**51%**	20%				0%
Avoided Loss/Gain (n = 33)	0%		9%	33%	33%	**21%**			3%
Foregone Gain/Gain (n = 35)	3%		3%		46%	**40%**		9%	0%

[a]: As the responses for several of the scenarios were obtained from four different questionnaires, Kruskal-Wallis ANOVA was employed to test for any between-group differences. All of the comparisons were nonsignificant for both the adjusted and unadjusted data set.

TABLE 2b: Adjusted' Frequencies - Experiment 2 (Losses)

	0	2/8	3/7	4/6	5/5	6/4	7/3	8/2	10/0
Q1	0%			1%	**96%**	3%			0%
Q2									
Direct Cost (n = 56)	4%			**86%**	9%				2%
Gain (n = 48)	0%				10%	**85%**			4%
Q3									
Foregone Gain (n = 50)	2%			12%	**78%**				8%
Avoided Loss (n = 54)	0%				**61%**	35%			4%
Q4									
Foregone Gain/Direct Cost (n = 25	0%			**52%**	20%	16%	12%		0%
Avoided Loss/Direct Cost (n = 31)	0%	26%	7%	**58%**	10%				0%
Avoided Loss/Gain (n = 23)	0%		13%	48%	4%	**30%**			4%
Foregone Gain/Gain (n = 25)	4%		4%		24%	**56%**		12%	0%

Appendix 1: Questinnaire for Experiment 1

This Questionnaire Is Part Of A MBA Thesis And Your Co-operation Is Greatly Appreciated.

Your Participation Is Voluntary & You May Withdraw From The Completion Of This Questionnaire At Any Time.

Please complete the following section first:

Gender: F:___ M:___ Age: ___ yrs. Indicate which class this is: _ _ _ _ _ _ _.

Is English your first language: Yes / No (cricle one)

===

INSTRUCTIONS

- In another experiment, students were randomly paired, e.g. student A with student B, student C with student D, student E with student F, etc., and asked how they would divide $10 between themselves under various conditions ('cases'.)
- For each case, the experimenter provided the $10.
- The (paired) students had identical information for each case but did not interact face-to-face.
- Each student submitted a final offer per case to an independent, anonymous arbitrator.
- This arbitrator would make a final, binding choice if the (paired) students' offers did not match. That is, the arbitrator would pick one or the other of the non-matching offers for each pair of submissions.
- Assume you are the arbitrator. For each of the cases below indicate how you think the students should divide the $10.
- Treat each case independently.

===

Case 1:

- Student A and student B were given $10 to divide between themselves.
- No costs were incurred by either student.

As the arbitrator, how do you think A and B should divide the money? **Circle one of the choices below.**

1 Each student should receive $5.
2 Student A should receive $10 and student B should receive $0.
3 Student A should receive $0 and student B should receive $10.

===

Case 2:

- Student C and student D were given $10 to divide between themselves.
- Student C had to pay $2 to participate in this activity.
- No cost outlay was incurred by student D.

As the arbitrator, how do you think C and D should divide the money? **Circle one of the choices below.**

1 Each student should receive $5.
2 Student C should receive $6 and student D should receive $4. [To account for C's $2 cost.]
3 Student C should receive $10 and student D should receive $0.
4 Student C should receive $0 and student D should receive $10.

PLEASE TURN OVER!

Case 3:

- Student E and student F were given $10 to divide between themselves.
- Student E had to choose between this and another risky activity. If s/he had chosen the other activity, s/he would have been paid $2 for sure.
- Student F did not have this choice.

As the arbitrator, how do you think C and D should divide the money? **Circle one of the choices below.**

1 Each student should receive $5.
2 Student E should receive $6 and student F should receive $4. [To account for E's $2 foregone gain.]
3 Student E should receive $10 and student F should receive $0.
4 Student E should receive $0 and student F should receive $10.

Case 4:

- Student G and student H were given $10 to divide between themselves.
- However student G had just won $6 in a lottery-type game preceding the experiment.
- Student H had to pay $2 to participate in the activity but did not gain any profits.

As the arbitrator, how do you think G and H should divide the money? **Circle one of the choices below.**

1 Each student should receive $5.
2 Student H should receive $6 and student G should receive $4. [To account for H's $2 cost.]
3 Student H should receive $8 and student G should receive $2. [To account for G's $6 gain.]
4 Student H should receive $9 and student G should receive $1. [To account for both H's $2 cost and G's $6 gain.]
5 Student H should receive $10 and student G should receive $0.
6 Student H should receive $0 and student G should receive $10.

Case 5:

- Student I and student J were given $10 to divide between themselves.
- Student J had the choice to participate in this and another risky activity. If s/he had chosen the other activity, s/he would have been paid $4 for sure.
- Student I had to pay $2 to participate in the experiment.

As the arbitrator, how do you think I and J should divide the money? **Circle one of the choices below.**

1 Each student should receive $5.
2 Student I should receive $6 and student J should receive $4. [To account for I's $2 cost.]
3 Student I should receive $3 and student J should receive $7. [To account for J's foregone gain.]
4 Student I should receive $4 and student J should receive $6. [To account for I's $2 cost and J's $4 foregone gain.]
5 Student I should receive $10 and student J should receive $0.
6 Student I should receive $0 and student J should receive $10.

- What do you think this questionnaire is about: ____________________

- Did your guess regarding the purpose of this questionnaire influence your choices: Yes / No. (Circle one.)

- If your guess influenced your choices, indicate how: ____________________

____

Thank You For Your Participation.

Appendix 2: Sample Questions (Q4) Used For Experiment 2

Case 4a:

- Student I and student J had to divide the $10 lottery loss between themselves.
- Student J had the choice between this lottery and another risky activity. If s/he had chosen the other activity, s/he would have been paid $4 for sure.
- Student I had to pay $2 to participate in the lottery.

As the arbitrator, how do you think I and J should have divided the loss? **Circle one of the choices below.**

1 Each student should pay $5.
2 Student I should pay $4 and student J should pay $6. [To account for I's $2 cost]
3 Student I should pay $7 and student J should pay $3. [To account for J's $4 foregone gain.]
4 Student I should pay $6 and student J should pay $4. [To account for I's $2 cost and J's $4 foregone gain.]
5 Student I should pay $10 and student J should pay $0.
6 Student I should pay $0 and student J should pay $10.

Case 4b:

- Student I and student J had to divide the $10 lottery loss between themselves.
- Student J had the choice between this lottery and another risky activity. If s/he had chosen the other activity, s/he would have been paid $4 for sure.
- Student I was paid $2 to participate in the task.

As the arbitrator, how do you think I and J should have divided the loss? **Circle one of the choices below.**

1 Each student should pay $5.
2 Student I should pay $6 and student J should pay $4. [To account for I's $2 gain]
3 Student I should pay $3 and student J should pay $7. [To account for J's $4 foregone gain.]
4 Student I should pay $8 and student J should pay $2. [To account for I's $2 gain and J's $4 foregone gain.]
5 Student I should pay $10 and student J should pay $0.
6 Student I should pay $0 and student J should pay $10.

Case 4c:

- Student I and student J had to divide the $10 lottery loss between themselves.
- Student J had the choice between this lottery and another risky activity. By choosing this lottery, s/he avoided losing $4 for sure.
- Student I had to pay $2 to participate in this lottery.

As the arbitrator, how do you think I and J should have divided the loss? **Circle one of the choices below.**

1 Each student should pay $5.
2 Student I should pay $4 and student J should pay $6. [To account for I's $2 cost]
3 Student I should pay $3 and student J should pay $7. [To account for J's $4 avoided loss.]
4 Student I should pay $2 and student J should pay $8. [To account for I's $2 cost and J's $4 avoided loss.]
5 Student I should pay $10 and student J should pay $0.
6 Student I should pay $0 and student J should pay $10.

Case 4d:

- Student I and student J had to divide the $10 lottery loss between themselves.
- Student J had the choice between this lottery and another risky activity. By choosing this lottery, s/he avoided losing $4 for sure.
- Student I was paid $2 to participate in this lottery.

As the arbitrator, how do you think I and J should have divided the loss? **Circle one of the choices below.**

1 Each student should pay $5.
2 Student I should pay $6 and student J should pay $4. [To account for I's $2 gain]
3 Student I should pay $3 and student J should pay $7. [To account for J's $4 avoided loss.]
4 Student I should pay $4 and student J should pay $6. [To account for I's $2 gain and J's $4 avoided loss.]
5 Student I should pay $10 and student J should pay $0.
6 Student I should pay $0 and student J should pay $10.

ATTACHMENT A

JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT 18, 227–237 (1990)

Environmental Policy Implications of Disparities between Willingness to Pay and Compensation Demanded Measures of Values*

JACK L. KNETSCH

Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

Received February 20, 1989; revised May 30, 1989

Like most economic analyses, environmental loss assessments and policy prescriptions are based on the assumption that, income effects aside, willingness-to-pay and compensation-demanded valuations will be equivalent. The accumulating evidence from controlled tests indicates, however, that the compensation measure is commonly far larger than the payment measure. Consequently, it is likely that, among other implications, losses are understated, standards are set at inappropriate levels, policy selections are biased, too many environmentally degrading activities are encouraged, and too few mitigation efforts are undertaken.

© 1990 Academic Press, Inc.

Environmental assessments and policy prescriptions, like most economic analyses, are commonly based on the empirical assertion that the willingness to pay and to accept compensation will yield equivalent measures of sacrifice. While budget restraints and wealth, or income, effects can give rise to small differences, it is assumed that "for many goods, services, and amenities that command a modest fraction of the consumers' budget, the differences between compensating and equivalent measures are trivial" [29, p. 244]. Consequently, the usual advice is that "As a practical matter it usually does not make much difference which of these two approaches . . . is adopted" [30, p. 125].

Given the presumption of equivalence, both losses and gains are usually assessed by the payment measure on the pragmatic grounds that "Generally speaking, willingness to pay is easier to estimate than required compensation" [20, p. 15]. Directives are clear: "The willingness-to-pay approach still holds: the true costs of unfavorable impacts are the total amount that people would be willing to pay to avoid them" [35, p. 152].

The accumulating evidence is, however, not at all consistent with these traditional assumptions of equivalence. The results of virtually all controlled tests of this conventional assertion, as well as the commonplace reactions of people to real choices, point instead to large differences between the alternative value measures. If this is the case, the implications for environmental assessments and policies appear to be substantial.

THE DISPARITY

The behavioral assertion of conventional theory is that people assess gains and losses by comparing the values of alternative end states—by, for example, compar-

---

*This research was in part supported by Fisheries and Oceans Canada and the Ontario Ministry of the Environment. Robin Gregory provided helpful comments on an earlier draft.

0095-0696/90 $3.00

Copyright © 1990 by Academic Press, Inc.
All rights of reproduction in any form reserved.

ing how they feel with their present wealth with how they would feel with their present holdings plus $10; or comparing their level of welfare with access to 10 parks and their well-being with 9 parks. However, instead of comparing alternative end states, people usually evaluate gains and losses in terms of changes from some reference position. And they value losses from this neutral point much more than they do gains beyond it [19]. Rather than the usual depiction of a smooth curve increasing at a declining rate, the evidence strongly suggests that people's actual behavior is characterized by a sharp break, or "kink," in the value function at the reference position.

This endowment effect and the resulting disparities in the valuation of gains and losses have been found in a wide variety of reported tests. Large differences were initially found in responses to survey questions about the values of a wide range of entitlements, including many assessments of environmental assets (see, for example, [15, 33, 13, 1, 3, 31, 4, 16, 14]). In all of these cases the minimum compensation people demanded to give up a good or an asset was reported to be several times larger than the maximum amount they were willing to pay to keep or acquire a commensurate entitlement.

Although the early studies were essentially unanimous in finding disparities ranging from three to over 10 times between the compensation and the payment measures of environmental values, these results have had virtually no impact on either policy prescriptions or economic analyses.

The earlier results did, however, prompt further tests of the valuation disparity using real exchange experiments. The results were the same. Even when exchanges of real goods and actual cash payments motivated the evaluations, the compensation demanded to give up an entitlement far exceeded the comparable payment measure of value [22, 16]. For example, in an initial test of valuation equivalence using real goods and actual cash payments, individuals were found to demand about four times more to give up a lottery ticket than they were willing to pay to acquire it [22].

While these results have been taken a bit more seriously (for example, [25 and 7]), they have had no more discernible impact on environmental policies or assessments than the earlier survey data. The findings from the real exchange experiments have been largely discounted on grounds of inconsistency with accepted economic theory or of being due to some unspecified artifact of the test designs. Others have suggested that some form of transaction costs probably accounts for the observed asymmetries, or that the disparities would likely be eliminated if experimental participants were allowed the opportunity to learn to be "rational" through repetition.

Some studies have shown that at least under some circumstances initial disparities between the two measures do decrease as participants gain experience from successive evaluation trials [6, 5]. However, these findings of a disciplinary impact of market experience were based on comparisons of the responses of buyers and sellers to either preliminary hypothetical questions or choices made in non-binding market trials, with their behavior in a subsequent binding trial with real monetary payoffs. Little, if any, further convergence was observed in these experiments once the trials became binding; that is, there was little evidence that learning led to equivalence over the range from the first to the last iteration of binding trials.

Also, these studies did not demonstrate that the discrepancy between buyers and sellers is eliminated after successive trials. Although the differences between the final selling and buying prices in an experiment involving the testing of a bitter substance (SOA) were not statistically significant, the ratio of median prices of sellers and buyers was still 2.6 to 1 [6].[1] In the study of the value of trees in a park, the ratio of mean bid and offer prices was indeed reduced by experience, from a high of 77 to 1 for the initial hypothetical survey responses, to 6.1 to 1 in the first potentially binding auction [5]. However, the disparity remained at 5.6 to 1 in the last binding auction—a final difference between valuations that seems to give little obvious support for the suggestion that this result shows that the discrepancy between measures "...can properly be eliminated by the forces of learning" [5, p. 554], or for the conclusion that valuations "...are more consistent with traditional economic notions of diminishing marginal utility" [5, p. 565].

A more recent series of experiments has provided further tests of the persistence of the disparity over repeated valuations and of the extent to which the observed differences are attributable to transaction costs [24]. The participants in these experiments first engaged in market exercises in which they traded assets which had values assigned to them by the experimenter.[2] These induced value markets were followed immediately by market exercises in which the same individuals traded real goods on the basis of values which the participants themselves ascribed to the goods in accord with their own preferences. The induced value markets and the subsequent goods markets were carried out with identical trading rules so that any transaction costs or similar inhibitions to giving up or acquiring the real goods should also cause reluctance to trade in the induced value markets. Any difference in outcome in the goods markets can thus be attributed to valuation differences. The markets for the real goods were repeated several times to allow for any learning from experience.

In each case the induced value markets and the markets for real goods yielded sharply different results. As expected, participants overwhelmingly made trades in the induced value markets in accord with the values that had been assigned. There was no inhibition due to transaction costs or lack of familiarity with the trading rules. As demonstrated in other studies, the induced value markets produced efficient and predictable outcomes.

In contrast, there was clear evidence of large and persistent valuation differences in each of the markets in which real goods were traded. Even though the trading institutions were identical to the preceding induced value markets and the participants had gained even more experience with the exchange rules, the outcomes invariably differed. Potential sellers of the real goods consistently placed significantly higher values on the goods than did potential buyers; differences generally ranged from two to six times more. Further, the disparate valuations of buyers and sellers remained stable over the repeated valuation iterations.

The differences in the outcomes of the induced value and the successive real goods markets are attributable to the systematic differences in the values that buyers and sellers placed on the assets being traded. In the induced markets, these values were given to each participant and they simply traded or did not trade in accord with these numbers. The endowment effect has a large influence on the valuations of real goods, however, and buyers systematically were willing to pay less for these assets than sellers demanded for the objectively commensurate entitlements.

---

[1] Further asymmetries in the experimental design and results are discussed by Knetsch and Sinden [23].

[2] Such induced value markets have been widely used to test various market institutions [34, 17].

A wide range of experimental designs have now been used to test the equivalence assumption and the persistence of observed disparities. This and other evidence indicates that, contrary to conventional assertions, losses are valued more than gains, the differences are pervasive and large, the disparities are not attributable to wealth effects, they are likely to persist over repeated valuations, and they are not the result of transaction costs or strategic behavior.

Further, the disparity and the difference in value that people usually attribute to opportunity costs and out-of-pocket costs seem to find expression in common references and normal practice. Examples include the frequent mention of "paper losses" to describe investment setbacks, the apparent different reactions to budget outlays and tax spending, the more stringent cost cutting by firms making losses than by those making profits, and the deliberate practice of framing prices as discounts or surcharges to gain more favorable customer reaction.

The available evidence strongly suggests that the conventional presumption of valuation equivalence may be a poor description of actual behavior and that its use may well lead to misleading assessments, inappropriate policy decisions, and resource misallocation.

IMPLICATIONS OF THE DISPARITY

The apparent dependence of valuations on entitlements raises questions about the propriety of various common economic analyses which are based on the presumption of value equivalence. For example, if the compensation demanded to give up a good is greater than the willingness to pay for like entitlements, as the evidence indicates, then indifference curves will not be reversible. Specific tests of the reversibility assumption have shown as much [21]. The assumption of transitivity may similarly be generally applicable for little other than transparent choices. Gains from trade may typically be overstated. And in much the same way the outcome, ascribed to the Coase theorem, that final allocations will be independent of initial entitlements in the absence of wealth effects and transaction costs will not result if values are higher for holders of an entitlement than for potential buyers. Experimental tests have shown that while bargains will result in efficient allocations, final outcomes will vary sharply with the initial assignment of rights [24].

Other implications of the valuation disparity are more closely linked to the economic analyses of environmental problems and the design of environmental policies.

Measurement of Losses

The valuation disparity findings suggest that most economic assessments of losses will be seriously understated if willingness to pay measures are used, and decisions will likely be biased because of this. It now does not seem to be the case that "...the circumstances in which this difference is important are probably quite extreme," and that therefore willingness to pay "should be the appropriate measure of benefits or losses" [32, p. 562]. The disparities at issue are not trivial ones attributable to wealth effects, but are those due to pervasive and large endowment effects.

Although most environmental management and policy decisions are not based on explicit estimates of economic costs and benefits, an awareness and appreciation of the economic consequences of alternative actions usually play a significant role in formulating options and choosing among them. It is common, for example, to base choices on some notion of what people seem to be willing to pay to obtain or keep an asset or entitlement, or the amount they might pay to avoid some form of environmental degradation. However, if environmental degradation or a contemplated change imposes losses on individuals, this usual practice of using the payment measure will likely lead to large understatements of the welfare changes. As a consequence, too many environmentally disruptive projects will be encouraged, too many harmful activities will be allowed, inadequate mitigation measures will be undertaken when environmental values are at risk, and compensation for losses will not fully indemnify adverse welfare changes.

Actions which appear to satisfy the potential Pareto efficiency criteria when losses are assessed by the payment measure may not do so when the compensation measure is used. Gains may be larger than the sums losers would be willing to pay to avoid the loss, but may be insufficient to match the amount that losers would require to agree to the change. The results of benefit–cost analyses may be biased in much the same way as costs will usually be underestimated by willingness to pay measures.

Another frequently omitted and otherwise understated cost is the loss imposed by the risk or uncertainty of the consequences of an action. For example, many of the physical impacts of waste disposal or damages due to pollutants are often unknown. And insofar as people are willing to sacrifice other things to be free of it, this uncertainty itself is a cost. If the uncertainties are imposed on people, these adverse changes "are then measured by the minimum compensation... that would fully indemnify their welfare losses" [27, p. 1214]. An accounting of the sums people are willing to pay to avoid the risk would likely greatly underestimate these costs.

Environmental Protection and Standards

Environmental control standards are, at least in principle, aimed at levels where marginal gains are matched by the marginal costs of further restraint. Levels based on willingness-to-pay measures of loss will likely differ from ones chosen on the basis of the compensation necessary for people to accept environmental degradation.

In an analogous way, standards of reasonable care may be misjudged and liability for damages may consequently be inappropriately assigned. Such decisions could reduce incentives for parties to undertake justifiable levels of protection from environmental and other harms. Again, more damaging activities and fewer preventive measures will be undertaken with the use of the willingness-to-pay measure of loss than would be encouraged with the usually more appropriate compensation measure.

Recognition of the greater weight given to losses than to gains may provide further justification for preserving environmentally sensitive areas and cultural and physical heritage resources. Frey and Pommerehne [11] have called attention to a collective endowment effect to explain a curious asymmetry in the international trade in art objects. Public, or collective, reactions to the possible loss of "national art treasures" are not only responsible for widespread legal restrictions on such sales, but give rise to far more successful fund raising activities when "a loss of the endorsement is to be prevented," than when "undertaken in order to buy some art object deemed worthwhile" [11, p. 474].

Much the same reaction may be prompted by the possible destruction of physical or cultural amenities or heritage resources. Efforts to establish national parks, for example, are usually not primarily motivated by the prospect of gaining a new amenity, but instead are focused on preventing losses of important physical assets. Similarly, the major incentive for designating heritage buildings is preventing their loss. This motivation suggests that the appropriate valuation of preservation activities is not how much people would be willing to pay to keep the resource in its present use or condition, but how much compensation they would demand to give it up.[3] In these cases, too, the choice of measure would not be a matter of indifference to the valuation.

The choice of the sustainability criterion that environmental resources should be maintained at levels sufficient to assure current productivity levels for future generations appears to be, at least in part, attributable to collective endowment effects [28]. Lowering sustainable production below current levels may be widely regarded as a loss, and both increased current consumption and a higher level of future productivity may be viewed as gains, thereby favoring a status quo standard over others that might be proposed.

Mitigation and Compensation

A further issue that may be affected by the valuation disparity findings is the relative preference for mitigation measures and compensation payments as alternative remedies for environmental harms or losses. Conventional economic analyses suggests that, all other things equal, compensation will normally be a more efficient and preferred means to deal with losses. The assumed advantage is due to the lack of restrictions attached to a compensation payment which permits recipients to use the funds for whatever good or service is of most value to them. An equal sum spent on mitigation would restrict reparation to the benefits of reducing the particular harm.

The results of the evaluation disparity studies suggest an alternative view of the relative merits of compensation and mitigation. People may view the compensation remedy as two events: a loss associated with the harm, and a gain of the money payment. As the compensation will be heavily discounted because it is viewed as a gain, more money will need to be paid to make up for any given harm. A mitigation measure, on the other hand, may well be treated as reducing the loss associated with the harm and will consequently be regarded as being more important.

The strength of this intuition was borne out by survey results indicating that affected parties may well value mitigation or replacement measures more highly than compensation payments, even when the size of the payments exceeds the expenditure on the mitigation works or when the mitigation seems to serve little beneficial purpose. For example, slightly over half of student respondents preferred to have an accidentally destroyed "old textbook used in an earlier course" replaced rather than receive cash that they could then use for another currently needed text or for any other purchase. Another sample expressed the same preference for replacement of a paperback novel which they had already read, even though this would appear to have less value than other things—including a new novel—that they could buy with a compensation payment.

---

[3] This does not necessarily mean that current owners must be paid to maintain such buildings. As noted below, the criterion for evaluation may not be dictated by existing legal entitlements to a cause of action.

From 80 to 90% of other respondents expressed a preference for spending money on reducing nuisances associated with a new waste disposal facility rather than paying compensation to people in the area, even when the payment was sufficient to make up for any loss in property values or was greater than the amount to be spent on the mitigating activities.

About 70% of another group of survey respondents thought that a new business should spend a large sum of money on efforts that would only be partially effective in overcoming a relatively minor environmental problem, rather than taking up the business's offer to spend a like sum "on whatever use is decided on by local residents in a referendum."[4]

Compensation policies do not appear to carry the efficiency benefits often ascribed to them on the basis of valuation symmetries. A more appropriate measurement of losses could lead to choices that would reduce costs and increase well-being.

Economic Incentive Schemes

Economic disincentive schemes in the form of effluent fees or pollution charges have long been advocated as an efficient pollution control strategy. But such proposals have generated strong opposition from many interests and have rarely been adopted.

The valuation disparity findings may offer a partial explanation for the relative unacceptability of the pollution fee strategies. The reactions to discharge fees and regulatory controls appear to be somewhat analogous to the ones that lead people to favour mitigation over compensation as a remedy for harms.

A control regulation would normally be framed as a measure that leads to a reduction of losses. On the other hand, individuals may treat the effluent tax as leading to two separate outcomes: allowing pollution, which they view as a loss; and payments made by waste dischargers, which they take as a gain. Because the two may not be integrated before valuation, most people may see the far more aversive loss as outweighing the benefits of the gain and therefore object to the proposal on grounds of perceived losses being larger than benefits.[5]

The transferable pollution permit proposal is likely a more attractive alternative to the pollution tax proposal. Despite their similar incentive and other characteristics, people may not have the same adverse reaction to the tradable permit option because, as for control regulations and not as for the pollution taxes, they will likely view the tradable permits scheme as leading to a reduction of losses.

---

[4] These results are also consistent with findings of community preferences in dealing with waste disposal site selection [36].

[5] In addition, people may also see those bearing the losses as not receiving any compensation and therefore suffering for the benefit of others—a violation of fairness norms that commonly decrease the acceptance of proposed changes [18].

A caveat to the alleged outcomes of transferable permits markets may, however, be in order. The normal presumption is that the initial distribution of permits is not a matter of much economic significance as market exchanges are assumed to assure that final allocations will completely reflect costs of abatement: "The [public] authority does not have to solve the cost-minimization problem to determine the initial allocation of permits: any allocation will do" [2, p. 187]. However, the final distribution of permits can be expected to be at least a partial function of the initial assignment of entitlements. Those receiving permits may be reluctant to give them up relative to what others will be prepared to pay to acquire them. Trades will doubtless take place, but asymmetries in valuations might be expected to lead to fewer transfers of permits than would be predicted on the basis of the traditional assumption of valuation equivalence. The initial distribution will therefore not be a matter of indifference to final outcomes or to total expenditures needed to attain any level of control.

This hesitation to give up pollution rights might be partially overcome by institutionally framing the nature of the exchange offerings in ways that reduce the aversion to the loss. The relative acceptability of permanent leases of art objects over sales of the same objects is an instructive example of the impact of such framing [11].

The valuation disparity findings are also inconsistent with the usual easy presumption of the symmetry of taxes and payments (or bribes) as means to control externalities: "A subsidy to reduce existing pollution (thus not available to any new pollution sources) has an allocative effect identical with that of the tax.... In both cases the opportunity cost of polluting is the same: the firm either foregoes a subsidy or pays an equal size tax for each unit of pollution it produces" [12, pp. 571, 572]. It now seems unlikely that giving up an opportunity cost of a foregone bribe will induce the same response as paying an out-of-pocket fee. It is more realistic to expect a charge to be framed as a loss and therefore provide a greater incentive effect than a subsidy which would be treated as a weaker motivating foregone gain.

Fairness, Licenses, and Compensation

The evidence that people value losses from a reference position much more than gains beyond it may well have implications for a wide array of quota allocations and reallocations, issuing and termination of licenses, and preferences regarding treatment of people experiencing adverse impacts of changes. Many of these are continuing sources of concern to environmental managers.

The asymmetric valuation of gains and losses appears to be a major reason that people judge some actions as fair and acceptable and others as unfair and unacceptable [18]. Changes or reallocations that are seen to impose a loss on one party for the benefit of another are very likely to be viewed as unfair and far less acceptable by a wide segment of the community that extends well beyond the individuals immediately involved.

The asymmetric valuation of gains and losses suggests that policies might be more acceptable to the extent that they are seen, for example, as measures that mitigate losses, as combining offsetting gains with losses to reduce feelings of loss due to the changes, as establishing a more visible link between current reductions and later gains to the same individuals, as breaking perceptions of gains to one group coming from losses imposed on another, and as framing losses in terms of foregone gains or as not necessarily being permanent.

Implementing new pollution controls that force current waste dischargers to reduce outflows may also widely be seen as imposing losses on polluters for the benefit of neighbors. This perception of losses being visited on particular individuals could give rise to a negative reaction requiring that the gains from controls be clearly demonstrated and be seen to greatly outweigh the costs before such changes will be judged to be worthwhile and command wide support.

The differing valuations of foregone gains and losses seem to influence the acceptability of changes in regulatory permit policies as well. For example, a reduction of allowable fish catch limits from historic levels is very likely to be viewed as a loss and to therefore have more severe adverse impact on economic welfare than denial of an increase in catch levels. And in those cases where a license is issued for a specific term—with no legal entitlement for any further extension—non-renewal may still be viewed as a loss from a reference position and would consequently be widely seen to be more unfair and less acceptable than denying a foregone gain of a similar right to others.

Reference Points and Loss Measures

The valuation disparity results raise the important question of what people regard as their reference point in judging gains and losses. Because reductions of losses weigh more heavily than commensurate gains, a control program to reduce pollution, for example, is likely to be far more valuable, and more likely to be worth doing, if it is regarded by people as reducing a harm than if it is taken to be an improvement.

A common presumption is that the choice of measure is entirely dependent on existing legal entitlements: "the value of a resource . . . will vary with the assignment of property rights" [26, p. 36]. It seems more likely, however, that the choice of measure is more appropriately a matter of the reference point that people actually use to judge changes. Moves that are seen as imposing costs or reducing losses will normally be best measured by the compensation measure; those that are viewed as improvements beyond the reference level would more appropriately be valued as gains by the payment measure.

Legal entitlements determine grounds for a cause of action: if damages or an injunction are or are not warranted. Such rights may or may not influence reference positions and the resulting choice of appropriate measure of welfare change. An example of this lack of determinancy of the legal regime was provided in a study of the impact of range laws in Shasta County, California [8]. One half of the county operated under open range law and the other half with closed range. In spite of the opposite assignments of liability for cattle damage to crops, and the fact that the laws were well known by residents in both parts of the county, cattle owners were held responsible in all cases. The reference position was consistent: crop damages caused by cattle were considered to be a loss despite the extant legal position.

CONCLUSIONS

The conventional assertion that values attached to gains and commensurate losses are equivalent, and the advice that "practically speaking, it does not appear to make much difference which definition is accepted" [10, p. 3], seem now to be incorrect for a large class of environmental and other values.

There is little or no evidence to support the traditional view of equivalence, and there is a large accumulation of new evidence to support the alternative view. The equivalence assumption may well be part of long held theories "that have been worked out in technical detail for nearly a century and are hardly controversial" [9, p. 21]. However, the continued reluctance to modify the assertion of symmetry appears to be increasingly costly. On present evidence, community welfare might be better served by environmental policies and allocations that take a more realistic account of the different values that people ascribe to gains and losses.

REFERENCES

1. N. Banford, J. L. Knetsch, and G. A. Mauser, Feasibility judgements and alternative measures of benefits and costs, *J. Bus. Admin.* **11**, 25–35 (1979).
2. W. J. Baumol and W. E. Oates, "The Theory of Environmental Policy," 2nd ed., Cambridge Univ. Press, Cambridge, UK (1988).
3. R. C. Bishop and T. A. Heberlein, Measuring values of extra-market goods: Are indirect measures biased?, *Amer. J. Agr. Econom.* **61**, 926–930 (1979).
4. D. S. Brookshire, A. Randall, and J. R. Stoll, Valuing increments and decrements in natural resource service flows, *Amer. J. Agr. Econom.* **62**, 478–488 (1980).
5. D. S. Brookshire and D. L. Coursey, Measuring the value of a public good: An empirical comparison of elicitation procedures, *Amer. Econom. Rev.* **77**, 554–566 (1987).
6. D. L. Coursey, J. J. Hovis, and W. D. Schulze, The disparity between willingness to accept and willingness to pay measures of value, *Quart. J. Econom.* **102**, 679–690 (1987).
7. R. G. Cummings, D. S. Brookshire, and W. D. Schulze, Eds., "Valuing Environmental Goods," Rowman & Allanheld, Totowa, NJ (1986).
8. R. C. Ellickson, A critique of economic and sociological theories of social control, *J. Legal Stud.* **16**, 67–99 (1987).
9. D. Feenberg and E. S. Mills, "Measuring the Benefits of Water Pollution Abatement," Academic Press, New York (1980).
10. A. M. Freeman, "The Benefits of Environmental Improvement," Resources For the Future, Washington, DC (1979).
11. B. S. Frey and W. W. Pommerehne, International trade in art: Attitudes and behaviour, *Riv. Int. Sci. Econom. Comm.* **34**, 465–486 (1987).
12. L. S. Friedman, "Microeconomic Policy Analysis," McGraw-Hill, New York (1984).
13. I. M. Gordon and J. L. Knetsch, Consumer's surplus measures and the evaluation of resources, *Land Econom.* **55**, 1–10 (1979).
14. R. Gregory, Interpreting measures of economic loss: Evidence from contingent valuation and experimental studies, *J. Environ. Econom. Management* **13**, 325–337 (1986).
15. J. Hammack and G. Brown, "Waterfowl and Wetlands: Toward Bioeconomic Analysis," Johns Hopkins Press, Baltimore, MD (1974).
16. T. A. Heberlein and R. C. Bishop, Assessing the validity of contingent valuation: Three field experiments, presented at the International Conference on Man's Role in Changing the Global Environment, Italy (1985).
17. E. Hoffman and M. L. Spitzer, The Coase theorem: Some experimental tests, *J. Law Econom.* **25**, 73–XXX (1982).
18. D. Kahneman, J. L. Knetsch, and R. Thaler, Fairness as a constraint on profit-seeking: Entitlements in the market, *Amer. Econom. Rev.* **76**, 728–741 (1986).

19. D. Kahneman and A. Tversky, Prospect theory: An analysis of decision under risk, *Econometrica* 47, 263–291 (1979).
20. A. V. Kneese, "Measuring the Benefits of Clean Air And Water," Resources For the Future, Washington, DC.
21. J. L. Knetsch, The endowment effect and evidence of non-reversible indifference curves, *Amer. Econom. Rev.*, 79, 1277–1284 (1989).
22. J. L. Knetsch and J. A. Sinden, Willingness to pay and compensation demanded: Experimental evidence of an unexpected disparity in measures of value, *Quart. J. Econom.* 99, 507–521 (1984).
23. J. L. Knetsch and J. A. Sinden, The persistence of the valuation disparity, *Quart. J. Econom.* 102, 691–695 (1987).
24. J. L. Knetsch, R. H. Thaler and D. Kahneman, Experimental tests of the endowment effect and the Coase theorem, working paper, Simon Fraser University (1988).
25. P. Knez, V. L. Smith, and A. W. Williams, Individual rationality, market rationality, and value estimation, *Amer. Econom. Rev.* 75, 397–402 (1985).
26. J. V. Krutilla and A. C. Fisher, "The Economics of Natural Environments," 2nd ed., Resources For the Future, Washington, DC (1985).
27. F. I. Michelman, Property, utility, and fairness: Comments on the ethical foundation of 'just compensation' law, *Harvard Law Rev.* 80, 1165–1258 (1967).
28. D. Pearce, E. Barbier, and A. Markandya, Sustainable development and cost benefit analysis, presented at the Workshop on Integrating Economic and Environmental Assessment, Vancouver, Canada (1988).
29. A. Randall, "Resource Economics," Wiley, New York (1987).
30. S. E. Rhoads, "The Economist's View of the World," Cambridge Univ. Press, Cambridge UK (1985).
31. R. D. Rowe, R. C. d'Arge, and D. S. Brookshire, An experiment on the economic value of visibility, *J. Environ. Econom. Management* 8, 1–19 (1980).
32. C. S. Russell, Publicly vested values for fish and wildlife: Comment, *Land Econom.* 58, 559–562 (1982).
33. W. F. Sinclair, "The Economic and Social Impact of Kemano II Hydroelectric Project on British Columbia's Fisheries Resources," Department of Fisheries and Oceans, Vancouver (1976).
34. V. L. Smith, Experimental Economics: Induced value theory, *Amer. Econom. Rev.* 66, 274–279 (1976).
35. E. Stokey and R. Zeckhauser, "A Primer for Policy Analysis," Norton, New York (1978).
36. C. Zeiss and J. Atwater, Waste facilities in residential communities: Impacts and acceptance, *J. Urban Planning Devel.* 113, 19–34 (1987).

ATTACHMENT I

ECONOMICS, LOSSES, FAIRNESS AND RESOURCE-USE CONFLICTS

*Jack L. Knetsch**

One major reason why it has proven difficult to accommodate conflicts ong users of natural resources is that in large part, voluntary economic rket exchanges are either not feasible or are an inefficient, inequitable, air, and otherwise inappropriate means of allocating among competing nands. Moreover, there is growing recognition that regulatory and other ı-market means of allocation often fail to take proper account of the full ıge of interests, frequently lack predictability, and are slow to respond to ınging values.

While present economic and legal institutions may appear to successfully .olve competing demands with a large measure of efficiency, equity, and :eptance in the majority of cases, there is also increasing awareness that :se institutions fail to provide satisfactory results in many specific, often ;hly visible and important, instances. And this is increasingly being .erpreted as a signal that the system may be systematically failing to provide cially desired allocations of entitlements in many other cases as well.

Economic issues are the source of many of the difficulties. Yet, many onomic analyses of resource conflicts have not made resolutions easier; in ırt, this may be due to their having provided misleading information.

The focus of this review is less on the better-known sources of allocation oblems, and more on recent findings of people's valuations of losses lative to gains, and on reasons why individuals find some changes fair and :ceptable and others unfair and unacceptable. The disparity between the nportance of gains and losses to people and the "rules" of acceptable actions ›pears both to add to the sense of unresolved conflict and to frustrate many ›nventional approaches to resolving conflict. A greater awareness and nderstanding of the gain/loss valuation differences and their impacts on ıirness judgments might, however, lead to more useful institutional reform.

FAILURES AND REJECTIONS OF MARKETS

Many natural resource commodities and services commonly trade in ıarkets and have economic values assessed in market prices, with supplies istributed accordingly. Timber, minerals, and fish are among those for which

commercial dealings provide useful investment incentives and allocation guidance.

Other natural resource outputs are not traded in markets or do so in sufficiently small degree that economic exchange yields little or no useful indication of social worth, and allocations may not correlate very well with wants and needs. Wilderness reservation, traditional hunting, outdoor recreation, and wildlife are prominent examples.

The specific causes of particular conflicts and allocation problems vary. Most, however, have in common a failure of legal and economic institutions to assure that all values, pecuniary and non-pecuniary alike, receive commensurate weight in allocation decisions or to provide any serious incentive for people to consider the possible harmful impacts of their decisions on other parties or on the long-term productivity of many resources. Efficient and acceptable allocations are often precluded, for example, by the lack of incentive to take account of external costs imposed on others, the motivation to overuse open-access common property, the lack of any private return to providing public goods, the distortions fostered by monopoly control, and outcomes that do not provide a satisfactory range of opportunities for various population groups to benefit from the use of resources.

When individuals, households, private businesses, statutory authorities, and government agencies responsible for environmental degradation are not required to bear these costs, they then do not pay the full costs of their activities, and have a perverse incentive to pursue activities that generate such losses and to impose them on others. The absence of an economic or legal need to fully weigh environmental or other such "third-party" costs contrasts sharply with the incentive to economize on the use of other goods or production inputs, and is a pervasive cause of resource conflicts.

Garrett Hardin characterized the misdirected incentives leading to environmental disruptions and depletion of unrestricted, or open, access common property as "the tragedy of the commons":[1]

> The tragedy of the commons develops in this way. Picture a pasture open to all. It is to be expected that each herdsman will try to keep as many cattle as possible on the commons. Such an arrangement may work reasonably satisfactorily for centuries because tribal wars, poaching and disease keep the numbers of both man and beast well below the carrying capacity of the land. Finally, however, comes the day of reckoning, that is, the day when the long-desired goal of social stability becomes a reality. At this point, the inherent logic of the commons remorselessly generates tragedy ....
>
> Adding together the component partial utilities, the rational herdsman concludes that the only sensible course for him to pursue is to add another animal to his herd. And

---

Much of this discussion is based on findings from research supported by Fisheries and Oceans Canada and the Ontario Ministry of the Environment, and the Social Sciences and Humanities Research Council of Canada.

[1] Garrett Hardin, "The Tragedy of the Commons" (1968) 162 Science 1243 at 1244.

> another .... But this is the conclusion reached by each and every rational herdsman sharing a commons. Therein is the tragedy. Each man is locked into a system that compels him to increase his herd without limit – in a world that is limited. Ruin is the destination toward which all men rush, each pursuing his own best interest in a society that believes in the freedom of the commons.

Destructive over-exploitation of fish stocks is but one illustration of conflict primarily due to the common-property nature of free access, which in this case gives little incentive for any one individual to conserve if this person knows that others can and will take any fish that he or she might leave for the purpose of maintaining future stocks.

Some commodities need not be consumed, or even seen, to be enjoyed and valued. Wilderness areas and habitat for rare species of wildlife, for example, derive much of their value from people's knowledge that they exist and have been set aside for long-term protection. Indeed, in some cases extensive use of the areas may degrade the resource and reduce its value. Given the inability to exclude non-contributors from benefiting from such resources and the incentive of beneficiaries to free-ride on the contributions of others, non-market provisions are usually necessary – provisions that often must come at the expense of uses with more conspicuous economic worth.

A further, and generally much less appreciated, cause of resource conflict and impediment to its resolution stems from people's varying perceptions of risks and their consequences. It is common to presume that people weigh only the size of the expected loss when considering the adverse impacts that may result from a risky choice – that is, that they frame possible losses solely in terms of the magnitude of any loss multiplied by the probability of its occurrence. However, recent research has found that people's aversion to possible losses varies greatly depending on the nature of the risk and the particular circumstances of their exposure. For example, even if all other things are equal, most people react far more negatively to a risk imposed on them by others than to one they assume voluntarily. They would also sacrifice much less to avoid a risk over which they feel they have some control compared to ones in which they feel a lack of influence; and they find risks that are more uncertain, are less familiar, and have delayed consequences much more aversive than ones of similar but familiar and more immediate impact.[2] Introduction of risks from chemical discharges from a new processing plant can, therefore, be fully expected to cause great anxiety and resistance from a resident population even if they can be shown that they have undertaken even greater risks to themselves by increasing use of their automobiles.

---

[2] Paul Slovic, "Perception of Risk" (1987) 236 Science 280.

For these and other reasons, many values associated with natural resources and resource uses do not find ready expression in market prices. The non-pecuniary nature of these values does not mean that they are not important contributors to economic welfare. Nor does it suggest that such resources or resource uses are any less economically valuable than other resources or goods that are provided and traded at market-established prices. The difference is only that a market is possible and desired in some cases, and is impossible or rejected in others. It is not that one class or the other is economically more valuable because of the circumstances under which it is made available and allocated.

While this continues to be the source of some confusion and continuing bias, market and non-market or non-pecuniary values are equally *economic* values. In both cases people are willing to give up other things and other resources in order to make them available or to prevent their destruction. It is this willingness to sacrifice that expresses or measures economic worth. The conversion of a plot of agricultural or timber land to a public park or a wildlife reserve, for example, may well increase the economic value of the resource even though the cash flow from the previous use of the land is not replaced by any cash or commercial transaction associated with the new use.

The non-pecuniary nature of many resource values – including the existence value or option value of having certain amenities available even if no present use is anticipated, the satisfaction of maintaining traditional entitlements, and freedom from fear of exposure to substances with uncertain consequences – does, however, make comparisons with other competing allocations more difficult.

ECONOMIC VALUES AND THEIR ASSESSMENT

The economic – pecuniary and non-pecuniary – importance of conflicting uses remains central to how resource use is allocated among competing claimants. However, given the uncertainties and lack of ready means of measurement, and the often-misunderstood nature of what is to be measured, many assessments have not necessarily raised the level of informed debate or improved proposed resolutions of resource conflicts. Quite apart from the obvious, and expected, self-serving exaggerations and biased assertions of interested parties, there are other more systematic worries about valuations and their use in judgments, allocation decisions, and regulatory reform.

The first is the fairly obvious discounting of less easily assessed non-pecuniary values in planning decisions, in litigation, and in reform proposals, relative to more easily appreciated – and more easily measured – financial data. Although there are numerous exceptions, courts, for example, still often treat some amenity values as "mere delights" and not the equal of commercial returns.

A second source of potential malassessment is the practice of taking rious expenditures, or cost outlays, associated with the use or development f a resource to represent the value of the resource itself — a feature of what om Power has aptly called "folk economics".[3] The usual accountings clude such analysts' employment-generating activities as estimating tourist xpenditures, studying local economic impacts, and using multipliers to etermine so-called spin-offs. Although these are frequently done with great chnical competence, this is not uniformly the case, and in nearly all cases ey reflect narrow interests of local boosterism rather than resource values.

The common attribution of associated spending and secondary or spin-off mpacts to the value of the resource prompting the expenditures leads to a onsensical conclusion that a recreation facility located in some remote area is ore valuable than an equally attractive one located close to users because it equires greater effort and cost to get to the more distant site. It is to say that mineral deposit deeper in the earth is more valuable than one closer to the urface because more wage income will be generated in bringing the ore up to he processing facility. Such accountings give rise to poor decisions and to the spectacle of officials solemnly attempting to construct an economic justification for preserving an important historical, cultural, or environmental site by calling attention to the largely irrelevant spending by a few visitors, while totally ignoring large real losses in economic welfare that would accompany demise of the asset.

A further, related problem with the counting of expenditure impacts is that if these associated benefits are to be credited as part of the gains attributable to a particular project, then the expenditures, jobs, and income that are not realized because of the spending that is displaced by this use must surely be debited as part of the costs. It is clearly misleading to count the secondary gains without counting the secondary costs. While there may be an interest in increasing or maintaining economic activity in a particular area, the justification for projects contributing to this turns on the added value of having employment and income accruing in one area rather than in another region; the justification of economic feasibility or efficiency is not demonstrated by such one-sided accounting. The practice of crediting spin-off gains and ignoring commensurate losses resulting from the displaced spending leads not only to a further over-supply of wasteful projects, but blunts the search for ones that will more nearly satisfy both objectives of more productive use of economic resources and maintenance of activity in areas of deserving concern.

---

[3] Thomas Michael Power, *The Economic Pursuit of Quality* (Armonk, N.Y.: M.E. Sharp, 1988).

The pervasiveness of using local impacts as measures of resource values is in part attributable to another difficulty in sorting out more reasonable responses to a wide array of policy issues: the asymmetry of interests and incentives between parties having a large but diffuse interest in the results and those having a smaller but concentrated interest. Small groups representing specific interests are often able to impose costs of uneconomic alternatives on another group because these costs are spread among a larger number of individuals. As the loss to any single member of the larger group may be relatively small, there is little incentive for any single person to be too offended, to become very interested in the issue, to protest, or to attempt to persuade others of the wastefulness of the proposal. The larger returns to each in the small group provide quite the opposite incentive.[4] This asymmetry often results in unbalanced information, a lack of critical review, and biased treatments of gains and losses.

There is also a too-easy assumption that market prices, which may adequately measure the value of a substitute, also measure the value to people who consider alternatives as being inferior to a particular holding to which they have a particular attachment. A compensation payment of the market price of another house, for example, is unlikely to leave a long-term resident as well off as before the taking of the person's home.[5]

GAINS AND LOSSES

A further asymmetry that may often seriously compromise efficient and equitable choices of entitlements is the commonly unappreciated disparity between measures of positive and negative changes in economic well-being.

Something has economic worth only to the extent that people are willing to give up something else in order to acquire or keep it. The usual textbook suggestion is, for example, "The economic value of something is how much someone is willing to pay for it or, if he [or she] has it already, how much money he [or she] demands to part with it."[6] There is little disagreement on this — welfare gains are correctly measured by the maximum sums that gainers would pay to obtain them and losses "are measured by the total number of dollars which prospective losers would insist on as the price of agreeing to adoption."[7]

---

[4] Randall Bartlett, *Economic Foundation of Political Power* (New York: Free Press, 1973).

[5] Jack L. Knetsch, *Property Rights and Compensation: Compulsory Acquisition and Other Losses* (Toronto: Butterworths, 1983).

[6] Richard A. Posner, *Economic Analysis of Law*, 3rd ed. (Boston: Little, Brown, 1986) at 11.

[7] Frank I. Michelman, "Property, Utility and Fairness: Comments on the Ethical Foundations of 'Just Compensation' Law" (1967) 80 Harvard L.R. 1165 at 1214.

However, rather than the compensation measure, a second measure of oss is commonly used to weigh adverse impacts and assign entitlements: the aximum willingness of an individual to pay to avoid giving something up. he reasons for the common use of this second, and usually inappropriate, easure, are:

1. that its calculation is often more convenient than the correct compensation measure, and
2. that it is widely assumed that for all practical purposes the two measures will result in fully equivalent assessments.

Any disparity between the willingness-to-pay (WTP) and the compensation-demanded (willingness to accept, or WTA) measures is confidently expected o be negligibly small and of no practical importance: "the differences between compensating and equivalent measures are trivial",[8] and "[a]s a practical matter it usually does not make much difference which of these two approaches ... is adopted."[9]

As a result, it has become common practice to base assessments for all manner of policy and planning purposes on this presumption of evaluation equivalence. No reckoning of any difference is made, or thought to be necessary. The possibility of differing values is generally ignored in proposals to resolve resource-use conflicts and deliberations leading to the assignment of property entitlements, as well as in benefit-cost analyses, assessments of damages, and the setting of health, safety, and environmental standards.

There is, however, little evidence to support this traditional view of equivalence. Instead, the available evidence – which has been accumulating for a dozen or so years – supports an alternative view of large and systematic disparities. People value losses much more than they do commensurate gains.

The recent research studies are consistent in finding that people require payments to give up entitlements that are from about two or three to well over five times larger than the maximum sums they would agree to pay to retain the same rights. The differences are not trivial, and attempts to understand and deal with resource conflicts will be severely hindered by continuing to ignore them.

A useful summary of the findings from the research on gains and losses is provided by the choices made by two comparable groups of individuals

---

[8] Alan Randall, *Resource Economics: An Economic Approach to Natural Resource and Environmental Policy* (Columbus, Ohio: Grid, 1987) at 244.

[9] Steven E. Rhoads, *The Economists' View of the World: Government, Markets, and Public Policy* (Cambridge: Cambridge University Press, 1985) at 125.

participating in a simple experiment.[10] Each individual in the first group was given a coffee mug and all were told that the mug was theirs to keep. They were then offered a four-hundred-gram Swiss chocolate bar of roughly comparable price in exchange for their mug.

The participants in the second group were offered the opposite choice. They were each initially given one of the large chocolate bars and then offered a mug in exchange for their chocolate bar. All other conditions were identical to those of the first group.

The participants in both groups had the identical opportunity of ending up with a coffee mug or a chocolate bar; there was no effort required to make any necessary trade and no uncertainty concerning the validity of the exchange offers.

Conventional economic assumptions about how people value gains and losses, which are the presumptions used in most economic analyses, including valuations of competing resource uses, offer a strong prediction of what should occur. The relative numbers of individuals preferring mugs to chocolate bars should be roughly equal in the two groups, as the relative preferences are not expected to be affected by which initial entitlement is awarded or by any difference in the value of gains and losses.

The results of the experiment showed a very different pattern of preferences from the one predicted by traditional economic principles. Rather than roughly equal proportions in each group preferring the chocolate bar and the mug, the numbers varied widely and systematically in accord with the direction of the necessary exchange.

Group	Prefer Mug	Prefer Chocolate
1. Give up mug for chocolate	89 percent	11 percent
2. Give up chocolate for mug	10 percent	90 percent

Nearly nine out of every ten (89 percent) of those individuals required to lose their mug in order to gain a chocolate bar (Group 1) indicated that they valued a mug more than the chocolate. Only one in ten favoured a loss of their chocolate bar to gain a mug (Group 2).

Participants' preferences were clearly not independent of the initial entitlement or the direction of trade offers, as is commonly assumed. The individuals taking part in this exercise valued a mug more than a chocolate bar when facing the prospect of losing the mug, but valued a mug less than e chocolate when given the choice of gaining a mug at the cost of giving up e chocolate.

---

[10] Further details are provided in Jack L. Knetsch, "The Endowment Effect and Evidence of Nonreversible Indifference Curves" (1989) 79 The American Economic Rev. 1277.

The reference position of people in the experiment mattered very much in rms of whether the good was valued as a potential loss or as giving up a ossible gain. In common with most experience, but in contrast with much of onventional economics, losses were worth more than fully commensurate ains.

The results of the simple coffee mug and chocolate bar exchange exercise re representative of a general finding from essentially all controlled tests of he equivalence proposition. An early example is the response of a sample of luck hunters in the United States who indicated that, on average, they would ay $247 each to maintain a wetland necessary to support present duck opulations, but would demand a minimum compensation of $1044 to agree o its demise — a bit over a four-fold difference.[11] In another early study, haler[12] found that people placed very different values on a .001 change in isk to their life depending on the direction of the change. The compensation hey demanded was one or two orders of magnitude greater to accept the risk a loss) than the largest amount they were prepared to pay to eliminate an dentical risk (a gain).

Similar disparities have now been reported by many investigators using a variety of methods to evaluate widely varied assets, and differences have been shown to be independent of transaction costs, repetition of trade offers, income effects, or wealth constraints.[13] The results of the many tests and studies are consistent in showing that losses matter much more to people than commensurate gains, and that reductions in losses are more valuable than forgone gains.

While these findings have until recently not proven popular with many economic practitioners — an unfortunate social cost of their reluctance to give up conventional assertions — they seem in accord with the strong intuitions of most people. Oliver Wendell Holmes summarized this intuition many years before the recent studies were undertaken in these terms:

> It is in the nature of man's mind. A thing which you have enjoyed and used as your own for a long time, whether property or an opinion, takes root in your being and cannot be torn away without your resenting the act and trying to defend yourself, however you came by it.

---

[11] Judd Hammack & Gardner Mallard Brown, *Waterfowl and Wetlands: Toward Bioeconomic Analysis* (Baltimore: Resources for the Future, 1974).

[12] Richard Thaler, "Toward a Positive Theory of Consumer Choice" (1980) 1 J. of Economic Behavior and Organization 39.

[13] Daniel Kahneman, Jack L. Knetsch & Richard Thaler, "Experimental Tests of the Endowment Effect and the Coase Theorem" (1990) 98 J. of Political Economy 1325.

> The law can ask no better justification than the deepest instincts of man.[14]

The empirical evidence seems to be consistent with Holmes' statement; it differs only in the finding that the reluctance to give something up does not necessarily occur only after "a long time." Once a reference position is perceived, people value gains and losses as departures from this neutral point, and motivations respond accordingly.[15]

Further, while individuals exhibit disparities between gain and loss values, or endowment effects as they are often called, it also appears that groups respond in similar fashion. Frey and Pommerehne note that collective endowment effects motivate many public efforts to protect groups against losses. An example is the restrictions many countries impose on the export of national art treasures, and the notable ease of raising funds to prevent the loss of such an object relative to the difficulty of securing support when "undertaken in order to buy some art object deemed worthwhile."[16]

If the valuation differences represent the general case, as the evidence now suggests, then people will react very differently to actions and policies than might be expected on the basis of the usual assumption of equivalence between gains and losses. The conventional assertion may prove to be a poor basis for predicting behaviour, and its continued use is likely giving rise to misleading assessments of changes in people's economic welfare and poor predictions of their acceptance of policy decisions.

FAIRNESS AND ACCEPTANCE OF ALLOCATIONS

The disparate valuations of gains and losses influences many choices, including people's judgments of which economic changes or policies are more acceptable to them and which are less so. Actions that impose losses on particular parties or groups are typically viewed as being more onerous and are consequently judged unfair.

Strong evidence reflecting such patterns has been found among both survey respondents and groups participating in real money exchange experiments.[17] Differing valuations of losses and forgone gains were clearly vident in the responses, for example, of two groups who were asked to judge he fairness of reducing workers' incomes in particular circumstances. One roup was told that the employer reduced wages by 10 percent. The other roup was told that the employer accomplished the same dollar reduction by liminating the workers' bonuses. Nearly two-thirds thought the wage cut nfair, but only 20 percent had a similar view of an equal loss in the form of oregoing a bonus. The wage cut was seen as a loss whereas the elimination f the bonus was viewed as a much less aversive forgone gain.

---

[14] Oliver W. Holmes, "The Path of the Law" (1897) 10 Harvard L. Rev. 457 at 477.

[15] Daniel Kahneman & Amos Tversky, "Prospect Theory: An Analysis of Decision Under Risk" (1979) 47 Econometrica 263.

[16] Bruno S. Frey & Werner W. Pommerehne, "International Trade in Art: Attitudes and Behaviour" (1987) 34 Rivista Internazionale di Scienze Economiche e Commerciali 465 at 474.

[17] See Daniel Kahneman, Jack L. Knetsch & Richard Thaler, "Fairness as a Constraint on Profit Seeking: Entitlements in the Market" (1986) 76 The American Economic Review 728, and Daniel Kahneman, Jack L. Knetsch & Richard H. Thaler, "Fairness and the

The primary "rule" that motivated most judgments was that it is unfair for one party to suffer a loss for the direct benefit of another party. Such "zero-sum" games were widely seen as unfair in all of the different contexts used, including employee-employer relations, landlord-tenant dealings, and buyer-seller exchanges.

For example, in responding to the following vignette, over 80 percent viewed the action of the store manager as unfair: "A hardware store has been selling snow shovels for fifteen dollars. The morning after a large snowstorm, the store raises the price to twenty." Even though the heavy snowfall made the shovels more desirable and valuable, raising the price was apparently seen by the overwhelming majority of people as imposing an unjustified loss on the consumer for the benefit of the store owner. The price change was, therefore, not considered acceptable.

Cutting wages to a current employee in the face of other workers willing to work for less was similarly viewed as one party's suffering a loss for the benefit of another, and was therefore thought to be unfair. And raising rents to a sitting tenant, even when other housing prices had gone up, was widely seen as unfair for the same reason.

Factors that divert the gain from the party initiating the change have a dramatic effect on the perceived fairness of the actions. If one of the parties changes, or if one party is not seen to gain at the other's expense, an action such as raising prices, which would normally be considered unfair, becomes acceptable to most people.

When respondents were asked, for instance, if they thought it was fair for a store to ration a single remaining toy by selling it to the highest bidder, three-quarters indicated that this was not acceptable. However, when the information that "the proceeds will go to UNICEF" was added, indicating that the gain would go to a charity rather than to the store owner, the same auction became acceptable to four out of five respondents. When the store was seen to benefit at the expense of the customer, it was judged to be unfair, but when the store was seen not to profit from the added outlay of the customer, then it became fair in the judgments of the respondents.

Similarly, even though raising rents to a sitting tenant was widely felt to be unfair, it became acceptable to impose the same increase on a new tenant. A new employee's receiving a lower wage is considered fair because there is little or no notion of a higher reference wage; but reducing the wage of a current employee imposes such a loss from the reference position of the former, higher wage, and is therefore thought to be unfair.

Raising prices to pass on costs is another important class of cases where one party is not seen to benefit from the imposition of losses on the other. Such actions are widely seen to be fair, even when losers are put to considerable inconvenience. In one case, respondents were asked about raising rents to "a tenant living on a fixed income" where an increase would mean "the tenant would have to move." However, because "the landlord's costs have increased substantially ... and the landlord raises the rent to cover the cost increases", the increased rent was viewed as fair by a ratio of three to one.

There seems to be great sympathy for allowing firms to maintain their reference level of return. Thus, passing on costs is widely seen as permissible quite independent of market conditions. And while cutting wages of current employees was judged unfair by most respondents, an exception was normally made in cases where the employer was said to be losing money. It then became acceptable to most people to cut even nominal wages. A limitation exists, however: that any losses must be seen to pertain directly to the transaction at hand. It was not considered fair for a landlord to raise rents to make up for the loss of another source of income.

Again, contrary to common views of economic behaviour, people appear to be strongly motivated by quite clear notions of fairness. Actions were judged to be fair or unfair in particular patterns, and people indicated a willingness to modify their behaviour accordingly, even at some cost to themselves and when there was little chance that others could retaliate and discipline their unfair behaviour.

LOSSES AND RESOLVING CONFLICTS

The apparent disparity between people's valuations of losses relative to gains and the related perceptions of what actions are likely to be considered as fair or unfair, do not come readily from standard views of economic behaviour. The disparity and fairness criteria are, however, likely to influence people's reactions to the behaviour of others and to public policies and actions, and will therefore have direct implications for the acceptance of new competing resource uses and the ways in which conflicts are resolved.

Assumptions of Economics" (1986) 59 J. of Business S285.

Although most policy decisions are not based on explicit estimates of economic costs and benefits, awareness and appreciation of the economic consequences of alternative actions usually play a significant role in formulating options and choosing among them. However, losses resulting from an action or policy are likely to be understated in deliberations based on the conventional view of losses being weighed equally with gains. Consequently, changes having actual net negative impacts will be unduly encouraged, damages will be under-assessed, compensation awards may be too small, inappropriate standards of protection against injuries will likely be set, inadequate mitigation measures will be undertaken, and choices of preferred legal entitlements will be biased.[18]

The extent to which disparities between valuation of gains and losses can influence policy preferences is illustrated by the choice between mitigation and cash compensation remedies. Convention suggests that, other things equal, it is always better, or more efficient, to choose monetary awards over mitigation. Such payments allow recipients to purchase whatever they value most, whereas mitigating works limit the remedy to whatever value is attached to reducing a particular harm, and this may be of less consequence to the affected parties. In other words, it is presumed to be better to give people money that they can use for food or holidays, or whatever they want, than to spend a like sum on restoring habitat for a species they may care little about or which they consider to be in ample supply.

The findings of large disparities between gains and losses suggest an alternative view of the relative merits of the options. Mitigation may be treated by people as reducing a loss, and worth correspondingly more because of this; compensation may be viewed as leaving the aversive loss intact in return for a monetary gain, which, by being a gain, is worth correspondingly less.

The strength of this intuition has been borne out in survey studies in which respondents indicated, for example, that they greatly preferred having a large sum of money spent on reducing — but not eliminating — even a minor harm to having an equal or greater sum spent on an alternative of direct benefit to them. For example, they preferred having money spent reducing stream sediment that only brought back fish numbers by 5 percent to having the funds spent on local park and recreation facilities. Further, when the alternative to reducing the sediment was putting the damage award into a

general fund to be used, presumably, on whatever was most urgently needed by the community — allegedly the most efficient solution possible — it was viewed even more unfavourably.[19]

The evidence suggests that the form of remedial action probably has a large influence on its acceptability. It seems likely, for example, that replacing losses with close substitutes would be favoured over less closely related measures. Activities that promote notions of participation, and those that provide benefits for more members of the community who are seen to have suffered losses, are similarly likely to be better received.

More generally, an appreciation of the intuition that losses are far more aversive than forgone gains, and that people have quite definite notions of fairness, can lead to a better understanding of reactions and can suggest alternative means of dealing with resource conflicts, ways that will probably be more acceptable and effective.

Negotiating resolutions of conflicts will, for example, be more effective to the extent that parties are asked to forgo gains rather than accept a far more aversive loss. A reduced salary bonus is easier to accept than a wage cut; ending a discount is less aversive than raising prices; and sharing profits of a development may be preferred to paying increased costs.

In a similar way, conversion of an environmentally valuable asset to another use that would yield benefits to others might well be seen to be unfair even though it may be thought efficient. Like imposing a wage cut on a worker because another is willing to work for less, converting a local park to a valuable office block is probably seen as unfair. The gain would likely be seen as obtained directly as a result of losses to prior users of the resource. Also, compensation for pollution abatement activities may be popular, not because it adds to efficient production — it usually does not — but because a reduction of the costs imposed is seen to be fair. And imposing a loss is regarded as far more unfair than denial of an equivalent forgone gain. Thus, prohibitions that preclude potential profits are widely sanctioned as acceptable, whereas restrictions that impose out-of-pocket costs are as often agreed to be unfair. Findings of liability and compensation rules are often compatible with this distinction and enjoy a large measure of community support as a result.[20]

---

[18] See Duncan Kennedy, "Cost-Benefit Analysis of Entitlement: A Critique" (1981) 33 Stanford L. Rev. 387; and Jack L. Knetsch, "Environmental Policy Implications of Disparities Between Willingness to Pay and Compensation Demanded Measures of Values" (1990) 18 J. of Environmental Economics and Management 227.

[19] Jack L. Knetsch, *The Relative Values of Mitigation and Compensation Remedies* (Working Paper) (Vancouver: Simon Fraser University, 1991).

[20] See Frank H. Stephen & Eric Young, "An Economic Insight on the Judicial Control of Planning Authorities' Discretion" (1985) 7 Urban Law and Policy 133; and David Cohen & Jack L. Knetsch, "Judicial Choice and Disparities Between Measures of Economic Value" Univ. of Toronto Law J. (in press).

For many of the same reasons, disruptive changes that are seen to reduce losses are more acceptable than ones that offer less valued gains. A waste storage and treatment facility that will effectively rid a community of current accumulations of inadequately treated materials will be favoured over a more efficient one that will also deal with wastes generated in other jurisdictions.

Just as passing on costs was generally accepted regardless of market demand and supply conditions, policies that use fees to recover costs are likely to be more acceptable than ones that appear only to add to the returns of the taxing jurisdiction. And, contrary to the conventional dictates of public finance counsellors, collections of funds diverted to "general revenue" accounts will be much less acceptable than ones "earmarked" for a specific, announced purpose related to the levy.

Similarly, redistribution of resource earning or use entitlements may be less disruptive if it is designed to be less than permanent in nature, rather than shifting the resource from one party to the direct benefit of another. For example, while the international sale of national art treasures is widely resisted, their loan to other countries is not — even if the loans are widely known to be permanent.[21] Similarly, payments to equalization funds could be limited to periods when specified inequalities existed — even though it might be widely appreciated that such inequalities will almost certainly persist for very long periods, even exceeding present lifespans.

While transfers from one party to the benefit of another are generally judged harshly, the research findings suggest that perhaps a large portion of the usual resistance might be overcome with the use of an acceptable standard to induce a reference level from which changes are framed by people as gains or losses. For example, payments into an equalization fund could be made more acceptable if the funds were distributed to regions recognized as being below some norm. Communities below this reference position would be seen as being in a loss position, and the redistribution payments would then be taken as reducing losses rather than as a less acceptable bestowing of benefits. The contributors might then also see their payments somewhat less as losses, and somewhat more as reductions to their gains and as payments to a fund that could also benefit them in the future. The use of "poverty lines" in jurisdictions throughout the world is a notable example of the success of such a strategy even though the poverty levels are widely known to be little more than arbitrary designations.

Distinctions between gains and losses — and as to whether a positive change is taken as reduction of a loss or a gain, while a negative change is perceived as a loss or a forgone gain, which to a large extent determines the desirability of changes or the acceptability of a proposed resolution of conflicts — are not necessarily determined by either the *status quo* or extant legal entitlements. Reference positions may well be based instead on an expected or normal condition, which may differ from the present and from legally enforceable rights. Reducing current pollution levels, for example, may be viewed as a reduction of a loss from the referent of a clean environment and therefore worth more than when viewed as an improvement from present levels of degradation.

And, contrary to common presumptions, extant legal entitlements do not appear to be a useful predictor of people's reference levels or determinant of the appropriate basis to be used to measure losses. The monetary measure of a welfare loss associated with the harmful action of a neighbour remains the sum of money that would just maintain the welfare of the individual at the level existing prior to the loss. The existing property rights determine whether or not an injured party has a cause of action against the neighbour, and therefore can or cannot obtain injunctive relief or receive compensation for the damage. The assignment of rights may influence the extent of the welfare loss felt by the injured party — violations of known rights may well prompt greater outrage than lawful changes — but the basis for assessing the loss is not determined by the entitlement.[22]

CONCLUSIONS

There seems to be little empirical evidence to support the conventional view that people weigh gains and losses equally. Losses appear to be a much greater concern to people, but this is often not reflected in the usual assessments or in many policy proposals. Recent findings also suggest that perceptions of fairness and related acceptance of changes are related to similar patterns.

On present evidence, some policy choices appear unnecessarily costly and generate acrimony that might be avoided; that is, ignoring the differential weighting of the gains and losses is likely to exacerbate the difficulty of resolving conflicts in resource use and make reasonable choices less certain. Improvements that take greater account of actual behaviour and reactions may well provide greater opportunities for more productive negotiations and eventual resolutions.

---

[21] Frey & Pommerehne, *supra*, note 16.

[22] Knetsch, *supra*, note 18.

ATTACHMENT J

DRAFT sept 93

THE REFERENCE POINT AND MEASURES OF WELFARE CHANGE[1]

Jack L. Knetsch
Simon Fraser University

(Prepared for the 3rd Canadian Conference on Environmental and Resource Economics, University of Ottawa, 1-3 October 93.)

The choice of economic welfare measure can have a very substantial impact on assessments, and may in many cases determine the outcome of the inquiry. However, selections and dictates such as:

> "the willingness of individuals to pay to avoid the negative impacts of emissions",
>
> "individuals' willingness to pay for a reduced incidence of illness and adverse symptoms",
>
> "the willingness to pay to prevent future occurrences of similar accidents",
>
> "a compromise solution has been reached by some ... the measure should lie between willingness to pay and willingness to sell",
>
> "the willingness to pay format should be used instead of the compensation required because the former is the conservative choice",

suggest that, in practice, many are made on the basis of varying combinations of confusion, dogma, mischievous misrepresentation, misplaced certitude and indifferent conformity.

Two widely held assumptions appear to be major contributors to the resulting arbitrariness and inappropriateness: (1) the various measures will, for essentially all practical purposes, result in equivalent assessments; and (2) the appropriate choice of

---

[1] This research was in part supported by The Social Sciences and Humanities Research Council of Canada, The Ontario Ministry of the Environment, and The Science Council of British Columbia.

measure to use in particular circumstances is prescribed by existing legal entitlements.

Neither assertion appears to be correct. Instead of indicating equivalence, the empirical evidence shows large disparities between the measures. And rather than serving as a useful guide, references to extant rights will often result in incorrect choices.

A set of criteria is presented below, based on a different reading of people's preferences, which suggests an alternative to the suppositions which appear to influence most present assessments.

The Measures and the Disparities

Changes in economic welfare are measured by economic sacrifice. They can be assessed by the maximum willingness of individuals to pay to secure or avoid a change, or by the minimum compensation required to leave such individuals as well off as they are without a change.

Aside from differences due to possible income or wealth effects, which are generally agreed to be trivial for most evaluations, the usual presumption is that estimates of an individual's maximum willingness to pay to avoid a loss will, for all practical purposes, be equal to the minimum compensation demanded to accept it. As Henderson put it early on, "...we shall normally expect the results to be so close together that it would not matter which we choose" (1941, p. 121). This conventional view that for nearly all measurements any differences will be inconsequential remains the presumption of choice -- "as a practical matter it usually does not make much difference which of these two ... is adopted" (Rhoads, 1985, p.125).

While the traditional assumption of equivalence of the measures has long been accepted and continues to justify present valuation practice, it has little or no empirical support. Instead, the evidence from many controlled tests consistently shows that losses matter much more to people than gains, and reductions in losses are more valuable than foregone gains. (Kahneman, Knetsch and Thaler, 1991).

The valuation disparities have been reported by many investigators using a variety of methods to evaluate widely varied assets. They were first noted in survey responses to hypothetical questions. Later real exchange experiments provided confirming results. Many of these studies centred on people being given a commonly traded good, and then offered the opportunity to exchange it for another good or for money (e.g. Kahneman, et.al, 1990). Other designs included offers to either pay people to accept a loss or allow them to pay to avoid it (e.g., Boyce, et.al., 1992). Contrary to the usual assertion that valuations should be the same whether the choice is to gain or to give up an entitlement -- "since a receipt foregone of a given amount is the equivalent of a payment of the same amount" (Coase, 1960, p.7) -- the numbers vary widely and systematically with the reference point of the individuals. These differences have now been shown to be independent of transaction costs, repetition of trade offers, income effects or wealth constraints (Kahneman, Knetsch and Thaler, 1990). While the results of many of these tests have used between subject designs, and therefore limiting the results to apply to a representative individual, similar differences in the valuation of gains and losses by the same individuals have been demonstrated as well (Kachelmeier and Shehata, 1992).

People have also been shown to exhibit similar larger valuations of losses relative to gains in their actual behaviour in making ordinary real choices (Kahneman, et.al., 1991). Frey and Pommerehne (1987), for example, note that collective endowment effects clearly motivate many public efforts to protect groups against losses. This is exemplified by restrictions which countries impose on the export of national art treasures and the notable ease of raising funds to prevent the loss of such objects relative to the difficulty of securing similar support for acquiring ones of equal or greater artistic merit. The valuation disparity, and the consequent reluctance to sell at a loss, has also been evident in the greater volume of house sales when prices are rising, over the number when they are falling, and similar smaller volume of sales of securities that have declined in price relative to those for which prices have increased (Shefrin and Statman, 1985). Firms frequently are reluctant to divest themselves of plants and product lines, even though they would not consider buying these same assets, and stock prices often rise when they do give them up.

The evidence indicates that the difference between valuations of gains and losses is a pervasive characteristic of people's preferences. Individuals appear to commonly require minimum payments to give up entitlements that are several times larger than the maximum sums they would agree to pay to retain them.

One implication of the disparity, of particular interest in deriving welfare assessment criteria, is that indifference curves will not be independent of the initial entitlement and direction of exchange offers (Knetsch, 1989 and 1992). People will commonly demand more of a good A to give up another good B, than they are willing to pay of A to obtain B. This produces a discontinuity, or kink, in the indifference curve at the initial reference position.

Legal Entitlements and Welfare Measures

The widely shared view that any differences between the measures are expected to be very small and of little practical importance has no doubt contributed to the meagre attention to the criterion selection problem. This lack of concern or at least enthusiasm, is apparent, for example, in the suggestion that, "...both measurements can be provided to policy makers and the associated value judgment of which to use can be left in their hands" (Just, Hueth and Schmitz, 1982, p.94).

For the most part, though, the prescribed choice has been seen as a matter wholly dependent on assigned legal entitlements. Thus, if a legal entitlement is permissive of an activity, then any value resulting from restricting that activity should be measured in terms of how much victims would pay to be relieved of the injury; alternatively, if there is a legal prohibition against an action giving rise to harms, these welfare losses should then be measured by the sums that would make the injuries acceptable to those adversely affected.

Almost without exception, economists, and others, deem that the choice of evaluation measure "....will vary with the assignment of property rights" (Krutilla and Fisher, 1975, p. 36). Even more explicit and current, is the view that: "the use of [WTA] in damage assessment involves the implicit assumption that the public at large owns the resource in question ... the use of [WTP] in damage assessment involves the implicit assumption that the public at large is not entitled to the services of the resource foreclosed by the damages" (Kopp and Smith, 1993, p.342).

The near unanimity of using legal entitlements to select the appropriate measure of a loss seems to stem largely from discussions, particularly among economists, of the compensating variation (CV) and equivalent variation (EV) measures of welfare change associated with price changes. The usual assertion is that the CV takes the initial utility level as the reference point whatever the direction of the price change. This is then seen to "... carry with them implicitly the presumption that the individual has no right or claim to make purchases at the new set of prices. In contrast, the EV is the lump sum payment (tax) which would have to be given to (taken from) the individual to make him [or her] willing to forego a price decrease (increase). This explanation contains the presumption that the individual has a right to the new price set and must be compensated if the new price set is not to be attained" (Freeman, 1979, p.45).

Linking the assignment of entitlement to the choice of welfare measure is both unwarranted and misleading. Legal entitlements, or property rules, determine whether or not the injured party has a cause of action against a neighbour and therefore can or cannot obtain injunctive relief or receive compensation for the damage. The choice of measure is an issue having to do with selecting a welfare criterion which best assesses actual changes in welfare in particular cases.

Entitlements define what claims receive recognition by the community and determine the degree of protection offered . Assignments determine, for example, which party to a conflict will prevail over the other and the extent of protection from the harmful actions of others.

The entitlements that are accorded priority are those that are judged to be more consistent with the efficiency, equity, fairness and other justice goals of the community (Calabresi and Melamed, 1972). The extent of welfare loss associated with various injuries may well be an important determinant in the assignment of specific entitlements -- violations of known rights may well prompt greater outrage than lawful changes -- and rules presumably evolve to offer greater protection against more important losses. To this extent, property rights are related to welfare losses. However, this does not imply that the appropriate basis for measuring those losses is in any way determined by the existing rules of entitlement (Knetsch, 1983 and 1984).

Rather than furthering allocative efficiency, the practice of using the property rights criterion to justify a choice of measure may well compromise this aim. For example, a municipal council might lawfully carry out a project yielding benefits of $100 per household, net of losses to each for which householders would be willing to pay an average of $50 to avoid or would require $200 to accept. Although the authorities might enjoy a legal entitlement allowing the development, using the WTP criteria for assessing the loss would encourage a project that would yield benefits that are far less then the welfare costs that are imposed. Similarly, if a private party were assessed damages of only $50 in such cases it could only lead to too many accidents and developments inflicting greater costs than benefits on the community.

While the choice of measure is widely accepted to be a matter of extant property rights, this criterion now appears to be very much inferior to criteria that take greater account of people's preferences and include those of willingness to pay for gains and compensation demanded for losses.

The Reference Point and Measurement Criteria

Criteria for welfare measures that take greater account of people's views of gains and losses, rather than legal entitlements, can be illustrated with portions of indifference curves such as those in Figure 1. Here, wealth, or income,[2] is measured on the vertical axis and quantity of some good or entitlement, X, on the horizontal. The indifference curves are kinked at the reference point (such as point J) to reflect the greater value of losses over gains.

Five individual cases can be distinguished, depending on the direction of having more or less of the good X, the individual's reference level of X, and the present level of use of the good. All five take W2 as the initial level of wealth. The initial level of provision of the good is Q1 in the first three cases, and Q3 in the last two.

Case 1: A Gain. The first case is a gain from the present level of X from Q1 to Q2, with Q1 as the reference level. This is a gain relative to the reference, and the increase in welfare associated with this is given by the maximum sum the individual is willing to pay (WTP) for this increase in X (to remain, at a minimum, on the indifference curve corresponding to the initial point J). This WTP is given by the decrease in wealth W2 W1. An example of such a gain is the WTP to acquire access to a new park.

---

[2] For the present purpose either measure will suffice. However, in light of recent findings that people's time preferences vary widely depending on particular circumstances, the one being a simple capitalization of the other may be fairly unrealistic. But, for here this is another matter.

Case 2: A Loss. The second case is a reduction in X from Q1 to Q3, again with Q1 as the reference level. The welfare loss is given by the minimum compensation that the individual will accept (willingness-to-accept, or WTA) to maintain the present level of welfare (to remain on the same indifference curve as the initial point J). This is given by the sum W2 W4. This will normally be much larger than W2 W1, because of the disparity between the valuation of gains and losses which gives rise to the kink in the indifference curve. This case is illustrated by the WTA sum to accept destruction of an existing park.

Case 3: A Foregone Gain. In this case, the reference is taken to be Q3 and the present use is Q1. The change is a reduction from the present Q1 to Q3. Since the reference is Q3, the use at a level Q1 was clearly one of being in the gains. Consequently, the move to Q3 is one of giving up, or foregoing, the gain, and the appropriate welfare measure is that of a gain, namely, the WTP to keep the increase above the reference, W2 W1. An example might be the temporary use of a plot of land for recreation while the owner awaits materials or whatever to convert the land back to some commercial use that would again preclude the public benefit. It is important to distinguish this case from those in which expectations, or whatever, are sufficiently well-formed to result in the shift of the reference point to Q1, which would mean that the disappointment would then be treated as a loss, rather than a foregone gain, and the welfare loss would then be the much larger WTA.

Case 4: Recovery of an Accepted Loss. This case is really little different from the first case, except that the reference is at the lower level Q3, implying acceptance of this state even if there were some history of access to the higher level. The change is a positive move from the Q3 to Q1. Here the gain is measured by the WTP of W2 W1, the same criteria as the first case. There may be fewer examples of this case (assuming that it is a separate one), but one might be a car owner's WTP to repair damage from an accident caused entirely by the motorist's own carelessness.

Case 5: Reduction of a Loss. The remaining fifth case is a positive change from Q3 to Q1, but with the reference level at Q1. Given the reference of Q1, being at Q3 is a reduced level of X and is "being in the losses", and an increase in X is clearly a reduction of the loss. The welfare measure is, therefore, the compensation that would be necessary for the individual to forego this improvement and accept the loss -- the WTA measure. This is an important class of cases that would include the WTA needed to accept a higher risk of harm brought about by wastes discharged by a factory in the neighbourhood. This case is widely misconstrued by defining the change of eliminating the harm as being a gain, or benefit, for which the WTP would be the correct measure -- as in "the WTP for reductions in the risks of premature mortality due to diseases with environmental causes". This practice of treating reductions in losses as gains, clearly results in serious understatements of the value of such changes.

Welfare Measures and Defining the Reference Point

The criteria detailed above focuses attention on defining, or discerning, the reference position from which people evaluate positive changes as either gains or reductions of losses, and negative changes as either losses or foregone gains. For example, the empirically important case of distinguishing between having a positive move being defined as either a gain (case 1) or reduction of a loss (case 5) depends on what individuals view as the appropriate reference. In many cases, this position will be fairly obvious; for example, in the wide acceptance that being free of bodily harm is the reference whether or not a person is being physically assaulted. In others, however, it may not. And in some cases it might be expected that some people will use one reference and other individuals will use another.

The reference point may be taken to be the status quo in many cases, but it might well be taken to be some expected, or what is regarded as a normal, level in others. As suggested by Tversky and Kahneman:

> A diversity of factors determine the reference outcome in everyday life, the reference outcome is usually a state to which one has adapted; it is sometimes set by social norms and expectations; it sometimes corresponds to a level of aspiration, which may or may not be realistic (1981, p.456).

The choice appears to turn on whether people consider themselves to be at a position without access or expectation of access to the good, or at a position in which they enjoy either current use or access to the good or the expectation that such access has, for whatever reason, some element of normalcy associated with it. If the individual considers the position to be one of current use or a normal expectation of such, and if the use is then withdrawn, the welfare loss would be measured by the sum necessary to leave the person on the same indifference curve -- the compensation demanded measure. Similarly, causing an injury or imposing a harm is not likely to be considered a norm, and its mitigation or elimination is therefore likely better considered a reduction of a loss than a benefit.

An individual may well be enjoying use of a good, even if there is no legal entitlement to such use or if the assignment of rights is uncertain. If the party holding the legal entitlement chooses to exercise the right to prohibit further use of the good, the current user might well suffer a welfare loss which is again best measured by the compensation necessary to leave that person as well off as before such a change. Similarly, an adverse ruling that would clarify an ambiguous entitlement and prohibit further use of the good by an individual, or would bring about a change in the nature of the good that would eliminate the previous enjoyment, would also result in a loss that is likely to be most appropriately assessed by the compensation measure. Again, the lack of legal entitlement only means that an individual harmed by a change has no cause of action either to block the change or to claim compensation for the resulting loss. It may also mean that the welfare loss may be less than if a right had been taken. However, it does not mean that the loss is therefore measured by anything other than the sum necessary to retain the previous level of welfare.

Ellickson's study (1987) of the response to differing range laws in Shasta County, California, in the U.S., provides an instructive example of reference positions not varying with legal entitlements. In one half of the County open range was the legal rule, and closed range the rule in the other half. Extensive interviewing of ranchers and farmers in both parts of the County revealed, first, that all parties were aware of the rules, but, second, that in spite of the opposite assignments of liability for crop damages caused by cattle, people followed similar remedies. Cattle owners were held liable in all cases, regardless of the legal rule. The reference position appears to be crops undamaged by straying cattle, and such harms are therefore considered losses for which the cattle owners were deemed responsible.

The possibility that a reference point from which gains and losses are evaluated may be something other than the current level, or status quo, suggests that some improvements may be best considered as remedial, and measured in terms of reductions of harms rather than as gains.

A reduction in current pollution levels may be such a case. The status quo may be a heavily polluted environment, and the change in welfare that would result from reducing waste discharges might vary depending on whether the action is viewed by the beneficiaries as an improvement in current conditions, or as mitigation of a present injury. If the current level of pollution and environmental quality is taken as the reference point, then the benefits of the improvement would presumably be measured by the WTP for an improvement -- a more modest sum that normally accompanies a gain. Alternatively, clean air may be regarded as the norm or reference level, and the change in quality will be valued, not as an improvement from the current level, but as a reduction in the current loss -- a sum that would usually be far larger than the willingness to pay measure.

The decision and the choice of measure in such cases depends on which criteria most accurately reflects the actual welfare change. If some notion of a pre-existing or normal level is widely regarded as the reference point, then the benefits associated with actions that are essentially corrective measures might best be measured by the compensation demanded to forego the reduction rather than the willingness to pay to obtain the change.

A recent investigation of the role of past states in determining people's reference positions, is an important example of an empirical basis for selecting an appropriate welfare measure (Gregory, Lichtenstein, and MacGregor, 1993). Different past histories of a variety of environmental problems were presented to different groups of subjects. Half of the participants were led to view current proposals to deal with the problems as improvements over historic levels, and half to view them as attempting to restore matters to levels enjoyed in the past. The results were clear: "Presenting a proposed change as the restoration of a prior loss rather than as a gain from the status quo is shown in all five cases to have a significant effect on the evaluation of environmental and health policy decisions" (p. 200). Further such findings could materially improve selections of measures and discourage inappropriate damage awards and inferior policy choices.

Conclusion

Many evaluation studies of the welfare changes associated with various proposed actions or policies have been fairly casual in choosing between willingness-to-pay and compensation-demanded measures. For the most part, the selection of the basis for measurement has been left as a matter of convenience. As a result, willingness to pay measures have been used not only to value gains, which is fully in keeping with normal efficiency criterion, but also to assess losses and the benefits of initiatives designed to mitigate losses, which is not consistent with this standard.

The justification for this practice has been that no important difference between the alternative measures was expected. This indifference can no longer be justified, however, as the compensation demanded measure will usually greatly exceed the willingness to pay measure.

The choice of measure is not a matter having much to do with legal entitlements of what is or is not lawfully permitted. Legal rules address questions of a cause of action to protect against injury or to recover damages; they do not determine the most appropriate basis for measuring welfare changes.

The choice depends on which measure most accurately reflects actual welfare change in the particular circumstances of each case. In the event of a loss, this is the amount that would leave initial welfare levels intact; it is not the willingness to pay to avoid the adverse change. And in the event of restoring a loss, it is the compensation demanded to forego the restoration, not the willingness to pay to improve matters. Use of "WTP for reduced incidence of illness" will lead to too many illnesses; and knowing use of WTP rather than WTA to assess damages because "this is the conservative choice" will compromise the deterrence function of compensation awards distorting incentives to attain efficient levels of damage avoidance -- to, for example, attain the proper mix of drunken and sober tanker captains.

References

Boyce, R.R., T.C. Brown, G.H. McClelland, G.L. Peterson, and W.D. Schulze, (1992), "An Experimental Examination of Intrinsic Values as a Source of the WTA-WTP Disparity", The American Economic Review, 82:1366-73.

Calabresi, Guido, and A. Douglas Melamed, (1972), "Property Rules, Liability Rules, and Inalienability: One View of the Cathedral", Harvard Law Review, 85:1089.

Coase, R.N., (1960), "The Problem of Social Cost," Journal of Law and Economics, 3:1-44.

Ellickson, Robert C., (1987), "A Critique of Economic and Sociological Theories of Social Control", Journal of Legal Studies, 16:67-99.

Freeman, A. Myrick, (1979), The Benefits of Environmental Improvement, Baltimore: Johns Hopkins Press.

Frey, Bruno, and Werner W. Pommerehne, (1987), "International Trade in Art: Attitudes and Behaviour", Rivista Internaxionale di Scienze Economiche a Commerciali, 34:465-486.

Gregory, Robin, Sarah Lichtenstein, and Donald MacGregor, (1993), "The Role of Past States in Determining Reference Points for Policy Decisions", Organizational Behavior and Human Decision Processes, 55:195-206.

Henderson, A.M., (1941), "Consumer's Surplus and the Compensation Variation", Review of Economic Studies, 8:117.

Just, Richard E., Darrell L. Hueth and Andrew Schmitz, (1982), Applied Welfare Economics and Public Policy, Englewood Cliffs, N.J.: Prentice-Hall.

Kachelmeier, S.J. and M. Shehata, (1992), "Examining Risk Preferences Under High Monetary Incentives: Experimental Evidence from the People's Republic of China," The American Economic Review, 82:1120-1141

Kahneman, Daniel, Jack L. Knetsch and Richard Thaler, (1986), "Fairness as a Constraint on Profit Seeking: Entitlements in the Market", The American Economic Review, 76:728-741.

Kahneman, Daniel, Jack L. Knetsch and Richard Thaler, (1990), "Experimental Tests of the Endowment Effect and the Coase Theorem", Journal of Political Economy, 98:1325-48.

Kahneman, Daniel, Jack L. Knetsch and Richard H. Thaler, (1991), "The Endowment Effect, Loss Aversion, and Status Quo Bias", Journal of Economics Perspectives, 5:193-206.

Knetsch, Jack L., (1983), Property Rights and Compensation, Toronto: Butterworths.

Knetsch, Jack L., (1984), "Legal Rules and the Basis for Evaluating Economic Losses", International Review of Law and Economics, 4:5-13.

Knetsch, J.L., (1989), "The Endowment Effect and Evidence of Nonreversible Indifference Curves," The American Economic Review, 79:1277-84.

Knetsch, J.L., (1990), "Environmental Policy Implications of Disparities Between Willingness to Pay and Compensation Demanded Measures of Values," Journal of Environmental Economics and Management, 18:227-237.

Knetsch, J.L., (1992), "Preferences and Nonreversibility of Indifference Curves," Journal of Economic Behavior and Organization, 17:131-139.

Kopp, Raymond J., and V. Kerry Smith, (1993), "Glossary of Terms for Natural Resource Damage Assessment", in R.J. Kopp and V.K. Smith, eds., Valuing Natural Assets: The Economics of Natural Resource Damage Assessment, Washington, D.C.: Resources For the Future.

Krutilla, John V., and Anthony C. Fisher, (1975), The Economics of Natural Environments, Baltimore: Johns Hopkins Press.

Rhoads, S.E., (1985), The Economist's View of the World, Cambridge: Cambridge University Press.

Shefrin, Hersh, and Meir Statman, (1985), "The Disposition to Sell Winners Too Early and Ride Losers Too Long: Theory and Evidence", Journal of Finance, 40:777-790.

Tversky, Amos, and Daniel Kahneman, (1981), "The Framing of Decisions and the Psychology of Choice", Science, 211:453.

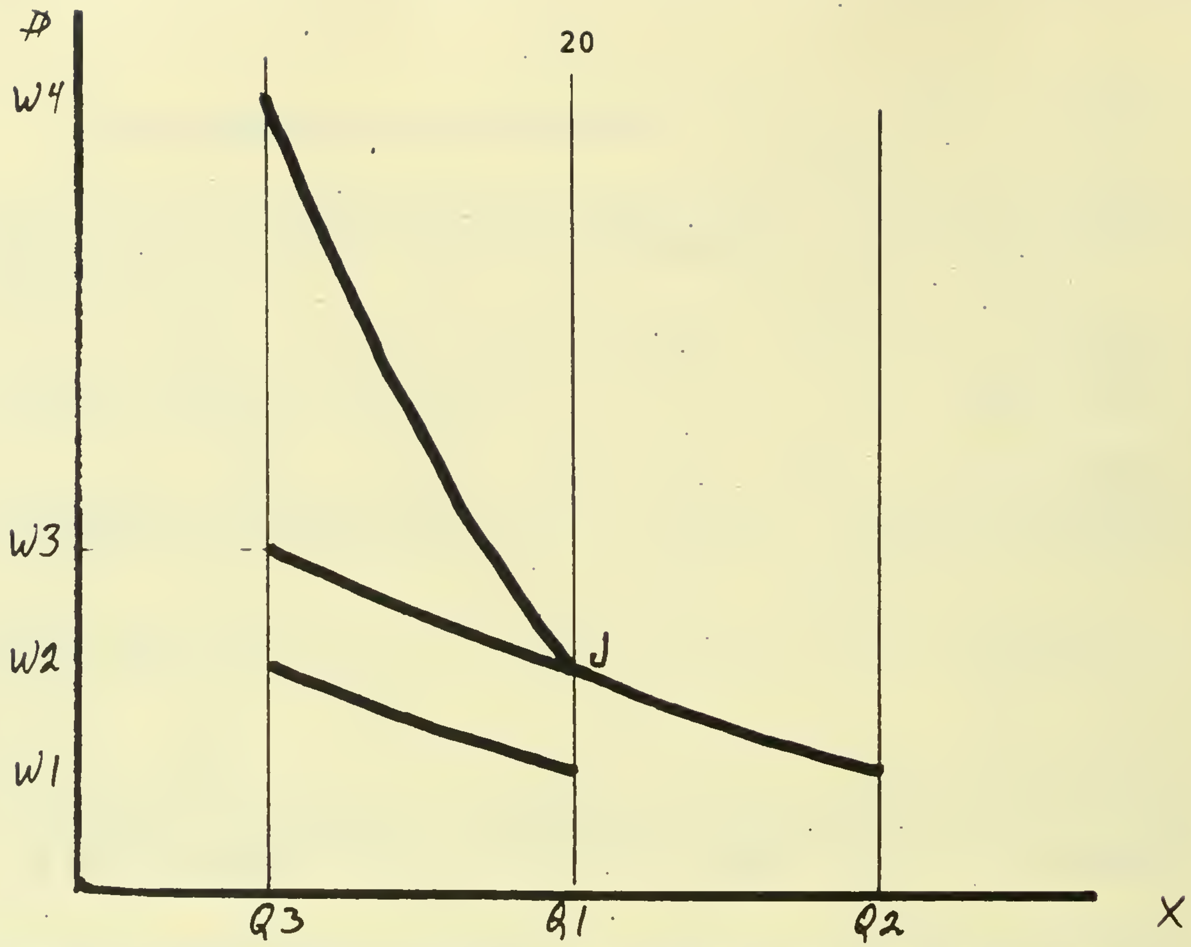


Figure 1. Portions of indifference curves depicting gains and losses, and criteria for welfare measures.

Present Level at Q1:

Reference Level at Q1:

Q1 to Q2 a gain measured by WTP of W2 W1 (Case 1)

Q1 to Q3 a loss measured by WTA of W2 W4 (Case 2)

Reverence Level at Q3:

Q1 to Q3 a foregone gain measured by WTA not to forego of W2 W1 (Case 3)

Present Level at Q3:

Reference Level at Q3:

Q3 to Q1 a gain measured by WTP of W2 W1 (Case 4)

Reference Level at Q1:

Q3 to Q1 a reduction of a loss measured by WTA to accept loss of W2 W4 (Case 5)

ATTACHMENT K

6Aug93

A TEST OF THE DOUBLE REFERENDUM CONTINGENT VALUATION QUESTION FORMAT

Daniel Kahneman and Jack L. Knetsch
University of California   Simon Fraser University

The dichotomous choice, or referendum, has emerged as the format of choice for contingent valuation questions. Asking respondents whether or not they favour paying a specified sum for a disignated purpose is believed to reflect the realism and simplicity of the common voter experience of casting ballots for or against a variety of expenditure proposals.

The ordinal ranking presented by the referendum choice of whether a respondent is willing or not willing to pay a specific sum is deemed to be an easier task than nominating a cardinal value of the maximum sum that the respondent is willing to pay, which is required in an open ended format. And asking a yes or no question of a respondent's willingness to pay[1] a specific single sum is thought to avoid anchoring problems inherent in sequential question formats in which the starting point of the sum posed in the initial question biases the final response.

A response to a referendum format question does, however, convey less information about the respondent's valuation than either an answer to an open-ended question or a final response of a sequential bid format. In these latter cases the respondent's maximum willingness to pay (or minimum WTA) is elicited. The referendum design involves asking individuals in different sub-samples whether or not they would pay a specific sum, with this amount varied among the

---

. The amount that a respondent would be willing to accept (WTA) could also e asked, though this is extemely rare even for monetary assessments of losse or which this is the more appropriate measure.

different groups. The mean (and median) willingness to pay is (essentially) calculated from the observed proportions of respondents saying that they would pay the different sums. Each respondent states only that he or she values the proposal more or less than a given amount, not how much more or how much less. Consequently, a larger sample is necessary to derive an equivalent assessment.

A double bounded dichotomous, or double referendum, format has been proposed, largely to overcome the lack of survey efficiency imposed by the limited information that can be drawn from single choice referendum questions (Hanemann, Loomis and Kanninen, 1991). In this variant a respondent first answers a yes or no valuation question, as in the single referendum, but is then asked a second question involving a willingness to pay a higher or lower sum depending on the answer given to the first question. Thus, one subsample of respondents might be asked, for example, if they would be willing to pay $100 for a specified purpose; those answering yes could be asked about their willingness to pay, say, $150, and those answering no could be asked about paying $50. In this way the answers reveal not just the proportion of the group willing to pay $100, the initial sum, but the proportions willing to pay more than or less than $50, and willing to pay $150 or more.

Given differing initial sums and answers from a second dichotomour choice question, a greater array of proportions of the total sample willing to pay varying sums is provided. Given all other things being equal, the double referendum format provides a more efficient survey design.

The double question design does, however, raise a concern that answers to the second question might be influenced by the sum posed in the first. The number of respondents willing to pay $100, for example, may be influenced by whether they had first been asked to pay, say, $20 or $200. This form of suggestion effect has been found to exert large impacts in many closely related exercises, and the existence of such anchoring would seriously compromise the legitamacy of the double referendum procedure.

A Test

The test was conducted in a series of telephone surveys in the Toronto, Canada metropolitan region, carried out by interviewers trained and supervised by a professional survey research firm. Equal numbers of female and male adult members of households were randomly sampled and the ten to fifteen minute interviews were held in evening hours.

The questions of interest appeared after some questions about the Canada Pension Plan and various consumer decisions. After asking: "Has anyone in your household gone fishing in the last Year?", the interviewer stated "we are interested in people's additudes to changes in the rivers and lakes of Ontario and their effects on fishing and recreation". Two WTP questions were then asked, as in the following example answered by one sub-sample:

> The number of fish in the lakes and streams of the Haliburton/Muskoka regions is expected to decline significantly in the future, because of natural causes and overfishing.

> Suppose that, unless fish breeding areas are continually protected and restocked, fishing will become poor even in areas where it is now considered good. Would you personally favour a $50 increase in your annual tax bill, to prevent this drop in fish population?

YES_____
IF YES, ASK:
Would you favour a $75 increase in your tax bill?
YES____ NO____

NO_____
IF NO, ASK:
Would you favour a $25 increase in your tax bill?
YES____ NO____

Respondents in each of the six sub-samples were asked variants of these questions which differed only in the proposed amounts they would pay. The specific sums posed in each of the six versions of the questionnair are detailed in Table 1.

Results and Discussion

The survey results are given in Table 2 and Figure 1. The valuations given on the basis of the initial willingness to pay question posed to respondents is given in the left column of Table 1, and corresponds to the single referendum format. These values are connected by the broken line in Figure 1.

The values given by the willingness to pay the amounts asked in the second questions are presented in columns to the right in Table 1. Consistency was assumed to the extent that the proportion of each sub-sample who said they would pay the lower sum (those saying no to the first question) were added to the proportion who had said they would pay a higher amount in the first question; similarly, respondents who refused to pay the amount in the first question were counted as unwilling to pay the higher sum asked of people who said they would pay the initial amount. The proportions willing to pay the three sums posed to each sub-sample are connected by solid lines in Figure 1.

A useful means of interpreting the results is to consider the configuration of Figure 1, on the basis of the (optimistic) null hypothesis that survey participants "read off" their answers to WTP questions from a stable system of preferences that are not subject to anchoring or other such effects. On the basis of this assumption all of the lines from each sub-sample should coincide for sums where they overlap. It is evident that they do not.

As measured by the responses to the initial question, the function that related acceptance to the amount drops relatively slowly, from 55 per cent willing to pay $25 to 21 per cent saying they would be willing to pay $300. In sharp contrast, the lines drawn to connect the responses from each sub-sample are relatively steep. This is particularly the case for the three sub-samples ~~in~~ for which the initial WTP question involved sums ranging from $25 to $100, ones common to many contingent valuation studies. The divergence between the within and between sample indications of the proportions willing to pay the various sums is evidence of a strong anchoring effect -- the responses of individuals to the second referendum question IS strongly influenced by the amount mentioned in the first.

The proportion of respondents stating they would be willing to pay $50 provides another indication of the anchoring effect. A WTP question for this amount was common to all of questionnairs. In one case the first question was for $25, a sum smaller than the $50; in another the first question was for the $50; and in all others the first questions asked about amounts greater than $50. Rather than roughly equal proportions of the randomly selected respondents in each sub-samples saying that they would be willing to pay $50, which would be expected in the absence of anchoring, the proportions varied widely and systematically in accord with a strong anchoring effect. Only 18 per cent of the respondents who were faced with $50 as an increase over the initial sum valued saving the fish by at least this amount. An intermediate 43 per cent indicated a value of $50 or more when this sum was asked in the first question. When $50 was asked following a request for a higher amount the proportion of the respondents who said the proposition was worth this much or more was as high as 63 per cent in one sub-sample, and averaged 59 per cent for all four of these sub-samples.[2]

The results of this test demonstrate a strong starting-point bias in the double referendum contingent valuation design. Two interpretations of this effect can be offered. The first focuses on the social rules that govern consecutive requests of the kind that occurs in the double dichotomous choice format. The anchoring effect could arise from two conventions that apply respectively to the behaviour of the respondent and the interviewer: (i) in normal interactions, refusing two requests in a row might be considered churlish, especially when the second request is much lower than the first; (ii) raising the ante after an affirmative answer to an initial request also might be considered intrusive and therefore likely

---

The data for these last four sub-samples suggest that a remote anchor ght be slightly less effective -- a result that would fit prior observation oted in the adaptation level literature (e.g. Helson, 1964).

to evoke negative reaction. Thus, the respondent may feel pressures to accept a more modest request after an initial refusal, and may be motivated to reject an escalating demand.

A more cognitive interpretation of anchoring is that the amount mentioned in the initial question provides information about the interviewer's expectations, thereby encouraging positive responses to this level but discouraging them for higher ones (Kahneman and Tversky, 1982). Whatever the nature of the cause, the results of this study suggest that the double referendum evokes an anchoring effect.

REFERENCES

Hanemann, Michael, John Loomis, and Barbara Kanninen, (1991), "Statistical Efficiency of Double-Bounded Dichotomous Choice Contingent Valuation", American Journal of Agricultural Economics, 73:1255-1263.

Kahneman, Daniel, and Amos Tversky, (1982), "On the Study of Statistical Intuitions", in D. Kahneman, P. Slovic and A. Tversky, eds., Judgments Under Uncertainty: Heuristics and Biases, Cambridge: Cambridge University Press.

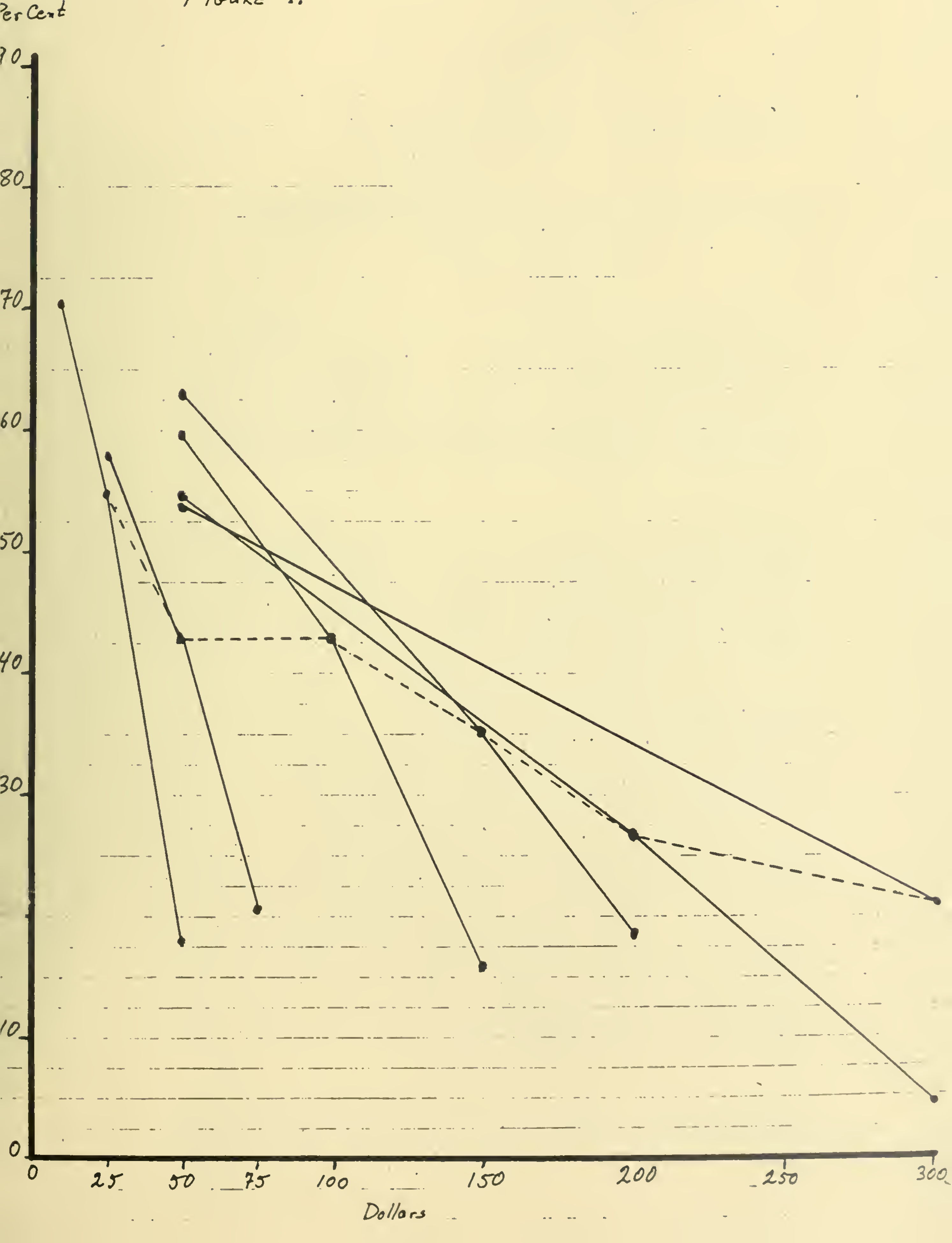
TABLE 1. Double Referendum Format Questions.

First Question WTP?	Second Question If Yes, Ask WTP?	 If No, Ask WTP?
$25	$50	$10
$50	$75	$25
$100	$150	$50
$150	$200	$50
$200	$300	$50
$300	$500	$50

TABLE 2. Proportion of Each Sub-Sample Willing to Pay Various Sums.

N	On First Question		On Second Question			
	WTP?	Per Cent	WTP?	Per Cent	WTP?	Per Cent
224	$25	54.9%	$50	18.3%	$10	70.6%
228	$50	42.9%	$75	20.6%	$25	57.8%
223	$100	43.0%	$150	16.6%	$50	59.6%
127	$150	35.4%	$200	18.9%	$50	63.0%
104	$200	26.9%	$300	4.8%	$50	54.8%
105	$300	21.0%	$500	8.6%	$50	54.3%

FIGURE 1.


Per Cent
80
70
60
50
40
30
10
0
0
25
50
75
100
150
200
250
300
Dollars

ATTACHMENT E

JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT 22, 57–70 (1992)

Valuing Public Goods: The Purchase of Moral Satisfaction*

DANIEL KAHNEMAN[†] AND JACK L. KNETSCH[‡]

[†] *Department of Psychology, University of California, Berkeley, Berkeley, California 94720; and*
[‡] *School of Resource and Enviromental Management, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6*

Received March 2, 1990; revised December 14, 1990

Contingent valuation surveys in which respondents state their willingness to pay (WTP) for public goods are coming into use in cost–benefit analyses and in litigation over environmental losses. The validity of the method is brought into question by several experimental observations. An embedding effect is demonstrated, in which WTP for a good varies depending on whether it is evaluated on its own or as part of a more inclusive category. The ordering of various public issues by WTP is predicted with significant accuracy by independent ratings of the moral satisfaction associated with contributions to these causes. Contingent valuation responses reflect the willingness to pay for the moral satisfaction of contributing to public goods, not the economic value of these goods. © 1992 Academic Press, Inc.

There is substantial demand for a practical technique for measuring the value of non-market goods. Measures of value are required for cost–benefit assessments of public goods, for the analysis of policies that affect the environment, and for realistic estimates of environmental damages resulting from human action, such as oil spills. In recent years the contingent valuation method (CVM) has gained prominence as the major technique for the assessment of the value of environmental amenities. This paper is concerned with a critique of CVM.

The idea of CVM is quite simple: respondents are asked to indicate their value for a public good, usually by specifying the maximum amount they would be willing to pay to obtain or to retain it. The total value of the good is estimated by multiplying the average willingness to pay (WTP) observed in the sample by the number of households in the relevant population. This value is sometimes divided into *use value* and *non-use value* by comparing the WTP of respondents who expect to enjoy the public good personally (e.g., benefit from improved visibility or from the increased number of fish in a cleaned up stream) to the WTP of respondents who have no such expectations. Specific questions are sometimes added to partition non-use value further into the value of retaining an option for future use, a bequest value, and a pure existence value [15].

The accuracy of the CVM is a matter of substantial practical import, not only in cost–benefit assessments but also in litigation over liability and damages. The validity of the technique is take as a rebuttable presumption in environmental cases brought in the United States under the Comprehensive Environmental Response, Compensation and Liability Act of 1980 (CERCLA). The research on the method has been reviewed in two authoritative volumes, which offer detailed

---

*This research was supported by Fisheries and Oceans Canada, the Ontario Ministry of the Environment, and the Sloan Foundation. Interviews and preliminary statistical analyses were performed by Campbell–Goodell Consultants, Vancouver, British Columbia. We benefited from conversations with George Akerlof, James Bieke, Brian Binger, Ralph d'Arge, Elizabeth Hoffman, Richard Thaler, and Frances van Loo, from a commentary by Glenn Harrison, and from the statistical expertise of Carol Nickerson.

0095-0696/92 $3.00
Copyright © 1992 by Academic Press, Inc.

guidelines for its use ([8, 17]; see also [11]). Some assessments of CVM have been very favorable, as illustrated by the claims that "the necessary structure for constructing a hypothetical market for the direct determination of economic values within the Hicksian consumers' surplus framework has been developed" [6, p. 173] and that contingent valuation "is potentially capable of directly measuring a broad range of economic benefits for a wide range of goods, including those not yet supplied, in a manner consistent with economic theory" [17, p. 295]. Acceptance of the technique is not universal, however, and some strong reservations about the adequacy of CVM to support specific compensation claims have recently been expressed [7, 9, 19].

The present article reports an experimental investigation of what is perhaps the most serious shortcoming of CVM: that the assessed value of a public good is demonstrably arbitrary, because willingness to pay for the same good can vary over a wide range depending on whether the good is assessed on its own or embedded as part of a more inclusive package. We next provide evidence for a similar difficulty in the response to payment schedules: WTP estimates can be much larger when the payment is described as a long-term commitment rather than as a one-time outlay. Another study suggests that WTP for public goods is best interpreted as the purchase of moral satisfaction, rather than as a measure of the value associated with a particular public good. Lastly, we examine the categories of expenditures from which contributions to public goods are drawn.

PART I. THE EMBEDDING EFFECT

The standard interpretation of CVM results is that the WTP for a good is a measure of the economic value associated with that good, which is fully comparable to values derived from market exchanges and on the basis of which allocative efficiency judgments can be made. However, two related observations that cast doubt on this interpretation have been discussed in the CVM literature. The first is an order effect in WTP responses when the values of several goods are elicited in succession: the same good elicits a higher WTP if it is first in the list rather than valued after others. For example, Tolley and Randall [22] found that estimates of the value of improved visibility in the Grand Canyon differed by a factor of three depending on whether this item appeared first or third in a survey. Because the order in which goods are mentioned in a survey is purely arbitrary, any effect of this variable raises questions about the validity of responses.

Another problem for CVM is an effect that we call *embedding*, also variously labeled as a part–whole effect, symbolic effect, or disaggregation effect [8, 17]: the same good is assigned a lower value if WTP for it is inferred from WTP for a more inclusive good rather than if the particular good is evaluated on its own. A finding that we obtained some time ago illustrates the embedding effect: the expressed willingness of Toronto residents to pay increased taxes to prevent the drop in fish populations in all Ontario lakes was only slightly higher than the willingness to pay to preserve the fish stocks in only a small area of the province (reported in [14]). It is quite unlikely that the respondents in Toronto viewed saving fish in the Muskoka area as a fully adequate substitute for saving fish in the whole province. The similar WTP observed for separate regions and for the whole of Ontario therefore appears anomalous. Further, the result raises a question about the proper assessment of WTP for a particular region: should this be estimated by the WTP assessment for that region in isolation, or by allocating to it a share of the sum offered for the cleanup of lakes in the entire province?

The embedding problem was noted long ago by investigators who were concerned with the appropriateness of aggregating WTP for several commodities obtained from different samples into an estimate of WTP for the package ([13], as summarized in [17, pp. 44–46]). Schulze *et al.* stated that "*no* researcher would be willing to defend the summation of CV values that have been obtained in various studies for *many* types of environmental effects; indeed the summation of average CV values for public goods thus far available in the literature would *exhaust* the budget of the average individual" [21, p. 6; emphases in the original].

The effects of order and embedding observed in assessments of value for public goods are difficult to reconcile with standard value theory. To appreciate why this is so, it is useful to consider the conditions under which assessments of the value of private goods would exhibit these effects. Two generic cases can be identified. The first involves goods which are perfect substitutes for one another and for which satiation is attained by the consumption of one unit. Thus, in the absence of opportunities for storage, resale, or altruistic giving, most adults will have zero WTP for a second large ice cream cone offered immediately after the consumption of the first. This is an order effect—the positive value of consuming an initial ice cream cone could be associated with any of those potentially available, but the value of any cone considered immediately thereafter would be zero. The value of ice cream cones under these circumstances also exhibits an embedding effect: WTP for 100 ice cream cones will not be higher than WTP for 1, much as WTP for improved fishing in all of Ontario was little more than WTP for fishing in a small area of that province. Although the notions of substitution and satiation may apply to some environmental goods, they do not readily extend to existence values for beautiful sites, historical landmarks, or endangered species. If it is found that WTP to save all threatened historical landmarks in a region is not much higher than WTP to save any single landmark, this can hardly be because each individual landmark provides as much utility as the whole set. Indeed, the uniqueness of the valued goods is the essence of existence value, as this notion has been discussed since Krutilla [15].

Effects of order and embedding are also expected in another extreme case: goods for which people are willing to pay a large part of their wealth. For example, the sum that an individual will pay to avoid the loss of both an arm and a leg is likely to be much less than the sum of WTP to save each limb separately, because the amount the person is willing to pay to save one limb is almost certain to be high in relation to available wealth, leaving little to prevent the second loss. In this case, order and embedding effects are produced by limited wealth. However, median WTP in CVM studies commonly falls in the range of $40–100 [9], far too small to be severely restrained by wealth.

The problem for the interpretation of CVM results is the following: if the value of a given landmark is much larger when it is evaluated on its own that when it is evaluated as part of a more inclusive package of public goods, which measure is the correct one? The discussions of the problem in the literature provide no agreed principles that would define the proper level of aggregation for the evaluation of a specific good. In the absence of such principles, the results of CVM become arbitrary. This criticism could be fatal. No measuring instrument can be taken seriously if its permitted range of applications yields drastically different measures of the same object.

Embedding a Public Good

Our first study was conducted to document the embedding effect in a controlled experimental design, focusing on the valuation of a public good that is of personal relevance to respondents: the increased availability of equipment and trained personnel for rescue operations in disasters. Coincidentally, the study was conducted within weeks of the San Francisco earthquake of 1989, a fact that certainly enhanced the relevance of the topic.

Three samples of adults living in the greater Vancouver region in Canada were interviewed by telephone. Samples were evenly split by gender. All calls were made in evening hours. The interviewers introduced themselves as being from a professional polling firm "conducting interviews on behalf of researchers at Simon Fraser University." All respondents were initially told:

> The federal and provincial governments provide a wide range of public services that include education, health, police protection, roads, and environmental services.

Respondents in one sample were then told to focus on environmental services, which were described as including "preserving wilderness areas, protecting wildlife, providing parks, preparing for disasters, controlling air pollution, insuring water quality, and routine treatment and disposal of industrial wastes." They were then asked the following question:

> If you could be sure that extra money collected would lead to significant improvements, what is the most you would be willing to pay each year through higher taxes, prices, or user fees, to go into a special fund to improve environmental services?

The evaluation questions were concluded at this point if the respondent's answer was zero. Other respondents were then asked:

> Keeping in mind the services just mentioned, including those related to providing parks, pollution control, preservation of wilderness and wildlife, and disposal of industrial wastes, I would like to ask you in particular about improved preparedness for disasters. What part of the total amount that you just mentioned for all environmental services do you think should go specifically to improve preparedness for disasters?

Subjects were allowed to answer by stating a dollar amount, a fraction, or a percentage. Where necessary, the interviewer immediately computed the dollar amount of the offered contribution and recorded that value. A third question was asked after some aspects of preparedness for disasters were listed (emergency services in hospitals; maintenance of large stocks of medical supplies, food, fuel, and communication equipment; ensuring the availability of equipment and trained personnel for rescue operations; and preparing for cleanup of oil, toxic chemicals, or radioactive materials):

> Keeping in mind all aspects of preparedness for disasters, what part of the total amount you allocated to improving preparedness do you think should go specifically to improve the availability of equipment and trained personnel for rescue operations?

TABLE 1

Willingness to Pay for Selected Classes of Goods and Allocations of Totals to Less Inclusive Groups

		Sub-sample		
Public good		Group 1 (N = 66) ($)	Group 2 (N = 78) ($)	Group 3 (N = 74) ($)
Environmental services	Mean	135.91		
	Median	50.00		
Improve disaster preparedness	Mean	29.06	151.60	
	Median	10.00	50.00	
Improve rescue equipment, personnel	Mean	14.12*	74.65**	122.64
	Median	1.00	16.00	25.00

*Two respondents did not answer this question, reducing N to 64.

**Four respondents did not answer this question, reducing N to 74.

The same procedure was followed with the second sample, except that the initial question they answered referred to "a special fund to improve preparedness for disasters," with a subsequent allocation to "go specifically to improve the availability of equipment and trained personnel for rescue operations." Respondents in the third sample were told to focus on preparedness for disasters and were asked to state their willingness to pay "into a special fund to improve availability of equipment and trained personnel for rescue operations."

Table 1 presents the medians and means of the willingness to pay responses for each of the questions in the three surveys. Zero responses are included in the calculations; respondents who stated a zero response to the initial question were assigned zero responses to subsequent allocation questions.[1] As in other applications of CVM, the data included extremely high responses, in some cases up to 25% of reported household incomes, which probably reflect a misunderstanding of instructions. These responses have considerable effect on the means of WTP, but there is no agreed way to draw a line beyond which responses will be rejected. To avoid this problem, our analyses of WTP results are based on medians, using all responses. The qualitative conclusions are unaffected by this choice of statistics.

The WTP for the public good mentioned in the first question posed to respondents was hardly affected by the inclusiveness of this good. The percentages of positive contributions were 61% for improvements in "the availability of equipment and trained personnel for rescue operations," 63% for "preparedness for disasters," and 65% for "all environmental services." The median WTP was $25 at the lowest level of inclusiveness and $50 at the two higher levels, but the difference was not significant in a Mann Whitney test in which the two higher levels were combined. The means of WTP across levels of embedding were also very close, and the differences among them did not approach statistical significance by F test. The pattern is the same as that which we observed in our previous study, in which WTP to preserve fish in all Ontario lakes was only slightly higher than WTP to maintain but a particular few.

---

[1] Glenn Harrison has suggested that this procedure could bias the results, because of the theoretical possibility that a respondent might be willing to pay for a good but not for a bundle that includes it. Note that this objection can be eliminated by informing respondents in advance that they will have an opportunity to allocate each contribution to an inclusive good among its separate constituents. It seems highly implausible that this minor procedural change would significantly alter results. The reasons for refusals to contribute in CVM surveys are commonly quite general (rejection of responsibility, opposition to extra taxes, etc.) and therefore likely to apply to the constituents as well as to more inclusive goods.

The bottom rows of Table I display the effect of position in the embedding structure on stated WTP for a particular good. The median amounts allocated to "equipment and trained personnel" vary from $25 when that good is evaluated on its own to $1.50 when the initial question concerns WTP for "environmental services." The three values shown in the last row of the table differ markedly from each other in the predicted direction. Group 1 differs significantly from Groups 2 and 3, by both parametric and nonparametric tests. The difference between Groups 2 and 3 only approaches significance.[2]

As in other studies, WTP values were small relative to reported incomes. Pooling over the three samples, the median WTP stated in response to the first question was $37.50 for respondents stating a family income under $20,000 (23% of total sample), $50 for income between $20,000 and $40,000 (39% of total), and $100 for families with incomes in excess of $40,000 (38% of total). The corresponding means were $97, $131, and $230. Clearly, these values are not in the range in which the embedding effect could be explained by constraints of wealth or income.

The results of this study demonstrate a large embedding effect. The key finding is that WTP is approximately constant for public goods that differ greatly in inclusiveness. The inevitable consequence of the insensitivity of WTP to inclusiveness is that estimates of WTP for the same particular good differ—by a factor of 16 for medians or 8 for means—depending on the scope of the initial question. An even larger embedding effect could probably be obtained by asking respondents to make explicit allocations to all the sub-categories at each level of embedding: the procedure of the present study, in which respondents make an allocation to a single subordinate good, appears likely to enhance its importance.

The specific good evaluated in the present study is fairly well defined, and the answer is interpretable as a quantity choice: how much extra equipment and personnel would you be willing to pay for? The good also has personal use value for most respondents, because improved availability of equipment and personnel for rescue operations would contribute to their safety and that of their families. The findings therefore extend the evidence for embedding: unlike demonstrations of embedding for existence value, the present results cannot be explained by invoking a concept of symbolic response [14; 17, p. 250].

Temporal Embedding of Payments

The embedding effects discussed so far apply to the specification of the good that is to be acquired. A related effect can arise in the specification of the schedule of payments. The question is whether respondents in a CVM survey are likely to make the appropriate discriminations between a one-time payment and a long-term commitment to a series of payments for a good. The issue is of some importance to the practical implementation of CVM. For example, willingness to pay was assessed in one study by asking people to state the amount that they would be willing to pay annually for 10 years in order to acquire a good, and conventional discount factors were applied to obtain an estimate of the present value of WTP [20].

---

[2] An earlier draft mistakenly stated that all three groups differed significantly from one another.

The issue of whether the participants in a CVM survey actually perform the discount calculations that are imputed to them has not been systematically examined, to our knowledge. The observations of embedding effects in which respondents did not discriminate between goods that vary in inclusiveness suggested the hypothesis that a similar failure of discrimination could be found between payments that vary in temporal inclusiveness. A small study was carried out to examine this hypothesis.

After completion of the main part of the interview, participants in the three groups of the first study were asked one of two versions of the following question. The version presented to each individual was selected at random:

> Now we would like to ask you how much you would be willing to pay (as a one-time payment/every year for a period of five years) to a fund to be used exclusively for a toxic waste treatment facility that would safely take care of all chemical and other toxic wastes in British Columbia.

Median WTP was $20, both in the group that considered a one-time payment ($N = 106$) and in the group that considered a five-year commitment ($N = 100$). The corresponding means were $141 and $81. The difference between the means was produced almost entirely by a few extremely high responses in the one-year group. There were five responses stating WTP of $1000 or more in that group, averaging $1800. There were only two such responses in the five-year group, averaging $1300. These extreme responses contribute approximately $54 to the difference of $60 between the means of the two groups. The results provide no reliable indication that the respondents discriminated between payment schedules that differed greatly in total present value.

The Significance of Embedding

What can be learned from these demonstrations of embedding? It may be useful to state the obvious qualification, that the present results have not established that insensitivity to inclusiveness is a *universal* characteristic of the valuations elicited in CVM studies. No single study could do so. The conservative conclusion from our findings is that future applications of CVM should incorporate an experimental control: the contingent valuation of any public good should routinely be supported by adequate evidence that the estimate is robust to manipulations of embedding, both in the definition of the good and in the specification of the number of payments. Whether this challenge can be met by appropriate CVM techniques is a question that will likely be the subject of further research.

Another defense of CVM against the embedding problem should be mentioned: the observation that different embeddings lead to different valuations of the same good would not be as troubling if there were a way of selecting one of these valuations as the correct one. As noted earlier, however, we were unable to identify in the existing CVM literature any compelling principles that could guide the choice of the appropriate embedding level for the good to be valued, or of a duration for the schedule of payments. Indeed, it is far from obvious that such principles can be found. Should the value of the damaged Alaska shoreline be assessed by WTP to clean up the damage done to it, by aggregating separate estimates of WTP to clean up parts of it, or by allocating to the cleanup a fraction of total WTP for environmental improvements? In the absence of agreed answers to such questions, our results suggest that current standards for the use of CVM may allow estimates of the value of a good that differ by more than an order of magnitude, all with an a priori equal claim to validity. As illustrated by the two examples reported here, the designer of a contingent valuation survey may be able to determine the estimated value of any good by the choice of a level of embedding. This potential for manipulation severely undermines the contingent valuation method.

PART II. THE PURCHASE OF MORAL SATISFACTION

The results presented in Part I do not support the interpretation of WTP for a public good as a measure of the economic value of this good. It remains a fact, however, that respondents express a willingness to contribute for the acquisition of many public goods, and there is no reason to doubt their sincerity or seriousness. Indeed, some elegant experiments have confirmed the willingness of people to pay for existence value—subjects actually paid to prevent a plant from being destroyed [5]. What is the good that respondents are willing to pay to acquire in such experiments or in CVM surveys? We offer the general hypothesis that responses to the CVM question express a willingness to acquire a sense of moral satisfaction (also known as a "warm glow of giving"; see [1, 2]) by a voluntary contribution to the provision of a public good. In attaining this satisfaction, the public good is a means to an end—the consumption is the sense of moral satisfaction associated with the contribution. An interesting feature of the warm glow of moral satisfaction is that it increases with the size of the contribution; for this unusual good, the expenditure is an essential aspect of consumption [18]. The interpretation of the responses to the hypothetical questions used in CVM in terms of moral satisfaction is consistent with Andreoni's economic analyses of actual donations to public goods, both in the field [1, 2] and in experimental situations [3], which distinguish the utility derived from increasing the total supply of the good from the utility gained in the act of giving.

Public goods differ in the degree of moral satisfaction that they provide to the individual making a contribution. Saving the panda may well be more satisfying for most people than saving an endangered insect and cancer research may be a better cause than research on gum disease. The quality of causes as sources of moral satisfaction will reflect individual tastes and community values. Our first hypothesis is that differences in WTP for various causes can be predicted from independent assessments of the moral satisfaction associated with these causes.

The results of Part I can be explained by invoking the additional hypothesis that moral satisfaction exhibits an embedding effect: the moral satisfaction associated with contributions to an inclusive cause extends with little loss to any significant subset of that cause. A closely related idea is that people may be willing to "dump their good cause account" on any valued cause [7]. Thus, contributing to the provision of rescue equipment may be as satisfying as contributing to the more inclusive cause of disaster preparedness. Indeed, a narrowly defined cause can be even more satisfying than a cause that includes it; it could be the case, for example, that saving the panda is more appealing than saving endangered species. Different subsets of a cause may vary in their appeal. In general, however, moral satisfaction could be expected to be about the same for an inclusive cause and for representative subsets of it.

An experiment was conducted to test these hypotheses. For the purpose of the experiment, a set of 14 pairs of public goods was constructed (see Table II). Each pair consisted of two causes, one of which was embedded in the other. The items were chosen to include two types of embedding: geographical embedding (e.g.,

TABLE II

Maximum Willingness to Pay for Various Causes and Ratings of Satisfaction from Making Contributions

Cause	Satisfaction Mean	WTP ($)	
		Median	Mean
Reduce acid rain damage in Muskoka, Ont.	7.18	20	40.91
Reduce acid rain damage in eastern Canada	7.25	50	214.55
Restore rural B.C. museums	4.67	10	32.78
Restore rural Canada museums and heritage buildings	5.79	20	113.47
Improve sport fish stocks in B.C. fresh water	5.25	10	41.89
Improve sport fish stocks in Canada fresh water	6.61	10	147.16
Protection for marmot, a small animal in B.C.	5.48	1.5	33.27
Protection for small animals in B.C.	6.42	10	141.75
Research on dengue fever, a tropical disease	4.57	0	52.42
Research on tropical diseases	4.97	4	17.83
Protection Peregrine falcon, an endangered bird	6.46	25	125.00
Protect endangered birds	6.98	20	59.07
Improve sport facilities in small communities in B.C.	6.22	10	209.75
Improve sport facilities in small communities in Canada	5.42	10	55.96
Rehabilitate recently released young offenders	5.78	50	233.16
Rehabilitate all recently released criminals	4.97	0	25.04
Habitat for muskrats, wild N. American rodent	4.70	0	51.60
Habitat for muskrats, squirrels, and other wild N. American rodents	4.59	4	52.28
Improve literacy of recent adult B.C. immigrants	6.30	10	190.53
Improve literacy of adults in B.C.	7.10	10	56.61
Replant trees in cutover areas in B.C.	7.80	20	151.70
Replant trees in cutover areas in western Canada	7.53	20	54.74
Increase research on toxic waste disposal	7.87	50	234.12
Increase research on environmental protection	7.44	50	98.77
Famine relief in Ethiopia	6.38	20	157.67
Famine relief in Africa	5.57	25	72.68
Research on breast cancer	8.12	50	243.14
Research on all forms of cancer	8.38	50	162.09

famine relief in Ethiopia or in Africa) and categorical embedding (e.g., research on breast cancer or research on all forms of cancer). The 14 issues were arbitrarily divided into two sets, labeled A and B, respectively, including the first 4 and the last 10 issues in Table II.

A special telephone survey of adult residents of the Vancouver region was conducted, with respondents randomly assigned into four groups. Respondents in groups 1 ($N = 60$) and 2 ($N = 61$) judged the moral satisfaction associated with the various causes (group 1 judged the inclusive items of set A and the embedded items of set B; group 2 judged the remaining items). After an introduction similar to that used in the surveys described in Part I and an indication that the questions were about "various causes to which people might be willing to make voluntary contributions," the instructions given to these two groups were:

> Please consider each of the causes separately and independently: that is, assume you are only being asked about the one cause. Indicate the degree of satisfaction you would receive from contributing to each cause on a scale from 0 to 10, with 0 indicating no satisfaction at all and 10 indicating a great deal of personal satisfaction.

Groups 3 ($N = 61$) and 4 ($N = 60$) were matched respectively to groups 1 and 2, but they provided measures of WTP for the same sets of causes. After the same introduction and indication of what the questions were about, these respondents received the following instructions:

> Please consider each of the causes separately and independently: that is, assume you are only being asked about the one cause. Indicate the most that you would be willing to pay for each.

The order in which the causes were presented was randomly determined separately for each respondent.

Table II presents the mean ratings of moral satisfaction and the medians and means of WTP for the 28 public goods included in the study, arranged in pairs. In each case, the embedded good is the first member of the pair. There is as usual a large discrepancy between mean and median WTP, due to extremely large WTP reported by a few individuals in each group. As before, we chose to focus on medians, without discarding any responses. The moral satisfaction ratings, which were made on a bounded scale, are not susceptible to large effects of a few aberrant responses, and the means of the satisfaction ratings were accordingly used in the analysis.

The hypothesis that WTP is predictable from assessments of moral satisfaction was tested by ranking the 14 issues evaluated by each group and by comparing the ranking of these issues by WTP and by moral satisfaction. The rank correlations between the means of satisfaction ratings made by group 1 and the median WTP of group 3 was 0.78 ($p < 0.01$). The corresponding correlation between the responses of groups 2 and 4 was 0.62 ($p < 0.02$). The general hypothesis that WTP can be predicted by ratings of moral satisfaction is strongly supported. As may be seen in Table II, there was only one striking discrepancy in the rankings of issues by WTP and by moral satisfaction: the rehabilitation of young offenders was one of the four causes eliciting the largest monetary contributions, but it ranked very low as a source of moral satisfaction. The discrepancy was not predicted, and any account of it must be speculative. One hypothesis is that the illegitimate context in which the need for public contributions arises makes it difficult to describe these contributions as yielding any kind of satisfaction—including moral satisfaction.

The data of Table II also allow a test of the effects of inclusiveness (embedding) on both moral satisfaction and WTP. The more inclusive causes have a very slight advantage on both dimensions overall, but the effect is weak and inconsistent: the more inclusive cause is associated with a higher rating of moral satisfaction for 8 of the 14 pairs of causes and with a lower rating for the other 6. Median WTP is higher for the inclusive than for the embedded cause in 6 pairs, identical in 6 others, and inferior in the remaining 2 pairs. On the other hand, mean WTP is higher for the embedded cause in 9 of the 14 pairs.

The results of this experiment support the proposed interpretation of willingness to pay for public goods as an expression of willingness to pay to acquire moral satisfaction. With only one salient exception, causes that were judged to provide little moral satisfaction also elicited relatively low WTP. Overall, there was a close correspondence between the rankings of issues by the two measures. Furthermore, the interpretation of WTP as an index of moral satisfaction helps explain the embedding effect: if the inclusiveness of the cause does little to enhance moral satisfaction, increasing inclusiveness should have little effect on WTP, as was indeed observed both in this study and in Part 1.

The Sources of Contributions to Public Goods

The question posed to respondents in a CVM survey is an unusual one, which has some features of a market survey, an opinion poll, and an appeal on behalf of a new charity. Respondents who follow instructions consider the possibility of a significant financial commitment to the provision of a public good. If they are serious about it, such a commitment to a new expenditure entails a corresponding reduction in other categories of spending. To understand the decisions of respondents in CVM surveys it is useful to identify the categories of spending from which they would expect to draw their contributions.

In the budget of most households there already exists a category of spending that is dedicated to obtaining moral satisfaction—voluntary contributions to charity. Spending on charity is far from negligible. For example, in fiscal 1988 donations by individuals in the United States totaled $86 billion, approximately $350 per capita. The respondent in a CVM study is likely to consider a new contribution in the context of the existing pattern of voluntary donations by the household. It is of interest to find out whether respondents view the proposed payment as a substitute to current charitable giving or as an addition to the moral satisfaction budget, requiring a reduction in other categories of spending. To test these possibilities, participants in the survey of Part 1 who had stated a positive WTP for the cause presented to them were asked a series of questions in the following format:

> Suppose you were actually called on to make the contribution to environmental services you indicated earlier. Which expenditure categories do you think this money would mainly come from? Would you spend less on ______?

The expenditure categories mentioned in the questions included food, charities, holidays (vacations), entertainment, savings, and "other things." After answering this series of questions, respondents who had listed more than one category were asked: "From which category do you think most of the money would come?" and the relevant list of categories was repeated to them. Table III presents the results.

TABLE III

Percentage of Respondents Indicating Reduced Spending in Various Expenditure Categories (N = 137)

Category	"Reduction"	"Greatest reduction"
Food	19.1	2.3
Charities	34.6	9.3
Holidays	64.2	15.5
Entertainment	76.1	41.1
Savings	46.2	22.5
Other things	65.9	9.3

The results indicate that added spending on environmental and disaster services would be drawn from discretionary spending, and especially from entertainment: Respondents would not expect to alter their eating habits. More important, they would not withdraw the contribution from current charitable giving. Most respondents apparently viewed the contribution as an addition to the "good cause" budget, not as a substitute for existing items in that spending category. The observed pattern of responses is the same as would be expected in answers to the question: "If you made an *extra* contribution to charity, where would it come from?" In the terms of the present analysis, the respondents appear to have considered the contingent valuation question as an opportunity to acquire additional moral satisfaction. Note that if households contribute only to causes that yield high moral satisfaction, the only way to increase the consumption of this good is by increasing contributions—to the currently favored causes or to equally satisfying causes.

CONCLUDING REMARKS

The research reported here had two objectives: to examine the proposition that CVM results are susceptible to an embedding effect that could render them largely arbitrary and consequently useless for practical purposes and to advance the interpretation of what people do in answering CVM questions.

The central result of the first study was that willingness to pay was almost the same for a narrowly defined good (rescue equipment and personnel) and for vastly more inclusive categories (all disaster preparedness, or even all environmental services). Correspondingly, the value assigned to the more specific good varied by an order of magnitude depending on the depth of its embedding in the category for which WTP was initially assessed. This result appears to invalidate a basic assumption of CVM: that standard value theory applies to the measures obtained by this method. As the choice of embedding structure is arbitrary, the estimates of value obtained from CVM surveys will be correspondingly arbitrary.

Our assessment of the validity of the CVM is in marked contrast to that reached by Mitchell and Carson [17] in their comprehensive review of the literature on this method. Mitchell and Carson recognized the potential severity of the embedding effect, but sounded a hopeful note in their discussion of it (p. 250), arguing that such an effect is not inevitable. The evidence cited for this conclusion was the observation that WTP to improve water quality nationwide, assessed in a national survey [16], was approximately twice as high as an estimate of WTP to raise the quality of water in the Monongahela River system in Pennsylvania, assessed in a local survey conducted by other investigators using a similar instrument [10]. Given the uncertainties of comparisons across studies and sampling areas, this evidence against embedding is not persuasive.

There was some prior reason to hope that the embedding effect might be restricted to non-use values, but the present results show that estimates of use value are not generally immune to embedding effects: an essentially complete embedding effect was obtained for disaster preparedness, a public good for which respondents have use value. Our tentative conclusion is that the factor that controls the magnitude of the embedding effect is not the distinction between public goods that have use value and those that only have non-use value. A more important distinction could be between public goods for which private purchase is conceivable and other goods for which it is not. Access to clean air and the right to fish in a stream could be privately purchased in a market, and sometimes are. The respondents in contingent valuation surveys have some experience in the purchase of such goods and could rely on this experience to determine their willingness to pay [6]. On the other hand, few respondents have experience in individual purchases of improvements in disaster preparedness, air traffic control, maintenance of species, or expansion of parks. The only way to procure such goods is by concerted public action, and the decision to make a voluntary contribution to such action has more in common with charity than with the purchase of consumption goods. Note that we do not assert that the CVM is necessarily valid for public goods that could be purchased by individuals. Our point is that the purchase of moral satisfaction is especially plausible as an interpretation of WTP for goods that could not be so purchased, even when these goods have use value.[3]

Students of CVM have long known that the respondents' answers to questions about their willingness to accept compensation for the loss of public goods (WTA) are strongly affected by moral considerations. Participants are prone to respond with indignation to questions about the compensation they would require to accept pollution of the Grand Canyon National Park, or of an unspoiled beach in a remote region. The indignation is expressed by the rejection of the offered transaction as illegitimate, or by absurdly high bids. The practitioners of contingent valuation hoped to avoid the difficulties of assessing WTA by substituting WTP even where WTA is the theoretically appropriate [8, 12]. The present results suggest that the adoption of the WTP measure does not really avoid moral concerns because the voluntary contribution to the provision of such goods can be morally satisfying. A treatment that interprets contributions to public goods as equivalent to purchases of consumption goods is inadequate when moral satisfaction is an important part of the welfare gain from the contribution [2]. The amount that individuals are willing to pay to acquire moral satisfaction should not be mistaken for a measure of the economic value of public goods.

---

[3] The application of CVM to goods such as hunting licenses in limited supply [4] is perhaps best viewed as a special case of market research, because these goods are in all essential respects conventional private goods.

REFERENCES

1. J. Andreoni, Giving with impure altruism: Applications to charity and ricardian equivalence. *J. Polit. Econom.* 97, 1447–1458 (1989).
2. J. Andreoni, Impure altruism and donations to public goods: A theory of warm glow giving. *Econom. J.*, 100 (June 1990).
3. J. Andreoni, "An experimental test of the public goods crowding out hypothesis." Working paper, University of Wisconsin (1990).
4. R. C. Bishop and T. A. Heberlein, Measuring values of extramarket goods: Are indirect measures biased? *Amer. J. Agr. Econom.* 61, 926–930 (1979).
5. R. R. Boyce, T. C. Brown, G. D. McClelland, G. L. Peterson, and W. D. Schulze, "An Experimental Examination of Intrinsic Environmental Values." Working Paper, University of Colorado (1990).
6. D. S. Brookshire, M. A. Thayer, W. D. Schulze, and R. C. d'Arge, Valuing public goods: A comparison of survey and hedonic approaches. *Amer. Econom. Rev.* 72, 165–176 (1982).
7. R. G. Cummings, Letter to Office of Environmental Project Review, Department of the Interior, November 10, 1989.
8. R. G. Cummings, D. S. Brookshire, and W. D. Schulze (Eds.), "Valuing Environmental Goods: An Assessment of the Contingent Valuation Method," Rowman and Allanheld, Totawa, NJ (1986).
9. R. C. d'Arge, A practical guide to economic valuation of the environment, *in* "Thirty-fourth Annual Rocky Mountain Mineral Law Institute Proceedings," Matthew Bendier and Co., New York (1989).
10. W. H. Desvousges, V. K. Smith, and M. P. McGivney, "A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality Improvements," Report to the U.S. Environmental Protection Agency, Washington, DC (1983).
11. B. Fischhoff and L. Furby, Measuring values: A conceptual framework for interpreting transactions with special reference to contingent valuation of visibility, *J. Risk Uncertainty* 1, 147–184 (1988).
12. A. M. Freeman, "The Benefits of Environmental Improvement: Theory and Practice," Johns Hopkins Press, Baltimore (1979).
13. J. P. Hoehn and A. Randall, "Aggregation and Disaggregation of Program Benefits in a Complex Policy Environment: A Theoretical Framework and Critique of Estimation Methods." Paper presented at the American Agricultural Economics Association Meetings, Logan, UT (1982).
14. D. Kahneman, Comments on the contingent valuation method, *in* "Valuing Environmental Goods" (R. G. Cummings, D. S. Brookshire, and W. D. Schulze, Eds.), Rowman and Allanheld, Totawa, NJ (1986).
15. J. V. Krutilla, Conservation reconsidered, *Amer. Econom. Rev.* 57, 787–796 (1967).
16. R. C. Mitchell and R. T. Carson, "A Contingent Valuation Estimate of National Freshwater Benefits," Report to the U.S. Environmental Protection Agency, Washington, DC (1984).
17. R. C. Mitchell and R. T. Carson, "Using Surveys to Value Public Goods: The Contingent Valuation Method," Resources For the Future, Washington, DC (1989).
18. H. Margolis, "Selfishness, Altruism and Rationality," Cambridge Univ. Press, New York (1982).
19. C. V. Phillips and R. J. Zeckhauser, Contingent valuation of damage to natural resources: How accurate? How appropriate? *Toxics Law Reporter*, 520–529 (1989).
20. R. D. Rowe, W. D. Schulze, and D. Hurd, "A Survey of Colorado Residents' Attitudes about Cleaning Up Hazardous Waste-Site Problems in Colorado," Report for the Colorado Attorney General's Office, Denver (1986).
21. W. D. Schulze, R. G. Cummings, and D. S. Brookshire, "Methods Development in Measuring Benefits of Environmental Improvements," Vol. II, Report to the U.S. Environmental Protection Agency, Washington, DC (1983).
22. George S. Tolley and A. Randall, "Establishing and Valuing the Effects of Improved Visibility in the Eastern United States," Report to the U.S. Environmental Protection Agency, Washington DC (1983).

ATTACHMENT M

JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT 22, 91-94 (1992)

REPLY

Contingent Valuation and the Value of Public Goods: Reply*

Daniel Kahneman[†] and Jack L. Knetsch[‡]

Department of Psychology, University of California, Berkeley, Berkeley, California 94720; and School of Resource and Environmental Management, Simon Fraser University, Burnaby, British Columbia, Canada V5A 1S6

Received March 26, 1991; revised April 15, 1991

We reported survey results in an experimental design that (i) showed little sensitivity of people's willingness to pay (WTP) to the inclusiveness of a public good and (ii) illustrated an embedding effect; WTP for a particular good varied greatly depending on whether it was assessed on its own or embedded in a more inclusive good [3]. The implications of embedding effects for the contingent valuation method (CVM) are potentially serious. Smith [4] does not deny—nor does anyone else to our knowledge—that a contingent valuation assessment that is demonstrably susceptible to manipulation by embedding is essentially meaningless as a measure of resource value.

Our study was intended to illustrate a simple test for the presence of embedding effects, which can be included in any individual contingent valuation study or added to it after the fact. Our demonstration was less elaborate than most contingent valuations, because it was designed as an experiment and because we had the freedom to pick a commodity that could be easily explained. The application of this design to more elaborate problems is straightforward. For example, it will usually be possible for a defendant in an environmental damages action to replicate a contingent valuation study in every detail, but also to add several alternative versions that systematically vary levels of embedding. If the various versions yield widely different assessments of value, and in the absence of a persuasive principle for choosing one over the others, any obtained estimate will be too arbitrary to provide either a basis for legal actions for damages or a useful guide to resource allocation policies. For example, if the estimated value of a group of sea otters varies by a sizable multiple depending on whether the assessments are derived from people's responses to questions about all otters, all sea mammals, or a single otter, none of the conflicting estimates can be used with confidence to assess the compensable loss for the destruction of any specific number of sea otters, or to determine the resources that should be devoted to avoiding such losses. This is the issue that our study raises.

Smith acknowledges the importance of the issue, but does not deal with it and devotes his article entirely to an unusually close scrutiny of the study itself. The

---

* We thank Ralph d'Arge, Baruch Fischhoff, George Loewenstein, and Richard Thaler for their comments and advice.

0095-0696/92 $3.00
Copyright © 1992 by Academic Press, Inc.
All rights of reproduction in any form reserved

reasoning appears to be that if sufficient doubts can be raised about this particular demonstration, the urgency of the issue will be somehow diminished. Probably because he did not find any fatal flaw in our argument, Smith offers a long list of points at which an alternative interpretation could perhaps be offered and a long list of alternative design choices that could perhaps alter the results. We find none of the alternative interpretations that are offered for our results compelling and are surprised by the extraordinary burden of proof that Smith would impose on a study that is critical of the CVM: the method is apparently to be judged innocent unless found guilty beyond any reasonable doubt, and the entire burden of removing all doubt lies on our study. This criterion is inappropriate to a debate in social science, where single studies are rarely conclusive. It is also worthy of note that Smith offers no definite prediction that the embedding effect will disappear if any particular improvement is adopted in further research. The refinements of method that he proposed are welcome as suggestions for future research on embedding, not as reasons to drop the subject. How Smith intended them is not clear.

To avoid both tedium and escalation, we restrict our response to Smith's discussion of the central point of our study, which is the demonstration of embedding. It is useful to recognize a distinction between two types of embedding effects: embedding can be called *perfect* if the same WTP is observed for nested commodities. To demonstrate *regular* embedding it is sufficient to show that WTP for a particular commodity varies depending on whether it is assessed on its own or as part of a more inclusive commodity. Perfect embedding, of course, implies regular embedding. We obtained perfect embedding in our study, but the argument against contingent valuation does not depend on this strong result. The weaker result of regular embedding suffices to label estimates of value as potentially arbitrary, and the weaker result is likely to be much more common. Thus, raising questions about our demonstration of perfect embedding does nothing to overcome the difficulty that regular embedding poses for the interpretation of contingent valuation. There are only two ways to defend CVM against a demonstration of regular embedding: (i) by offering improved survey techniques that reliably yield similar estimates of value for a commodity regardless of its level of embedding; (ii) by providing a rationale to select one of the competing estimates where they disagree. As far as we can tell, Smith did neither. Many of his observations are concerned with reasons for questioning the generality of the perfect embedding that we observed—but this is not the core of the problem.

The critique introduces our paper with a somewhat misleading description of its conclusions, to the effect that CVM-based estimates are "largely arbitrary and consequently useless for practical purposes." This phrase is taken from the statement that one of the objectives of our study was to "examine the proposition that CVM results are susceptible to an embedding effect that *could* render them largely arbitrary and consequently useless for practical purposes..." (emphasis added). Our key conclusion was clearly stated elsewhere. Following a comment noting that no single study could demonstrate that the embedding effect is universal, we wrote, "The conservative conclusion from our findings is that future applications of CVM should incorporate an experimental control: the contingent valuation of any public good should routinely be supported by adequate evidence that the estimate is robust to manipulations of embedding... ." We stand by this conclusion.

Section 2 compares the method of our survey to the criteria laid out by Fischhoff and Furby [2] and finds it wanting. The main conclusion is that "it is not clear that respondents understood the inclusive commodities . . . as KK intended." There is little force to this criticism. First, our experiment was designed to minimize the problem of ambiguity by using a commodity that is easily described. Second, the definition of "disaster preparedness" given to group 2 explicitly included "availability of equipment and trained personnel for rescue operations" as a constituent of preparedness – this removed the ambiguity about which Smith complains, but did not eliminate the embedding effect.

We do not claim that all our subjects understood the problem exactly as we do: no survey researcher can make that claim. Our point is that there is no reason to believe that ambiguity was particularly severe in our study. The ambiguity issue can be raised about any inclusive public commodity that is described without technical specifications and a detailed budget. Contingent valuation surveys are also notoriously vague about the critical issue of payment vehicle. It is inappropriate to hold a study that is critical of CVM to methodological standards that the targets of the critique do not, and probably could not meet: the Fischhoff–Furby criteria describe an unattainable ideal, not the current state of the art in contingent valuation (Fischhoff's review of the Mitchell–Carson book [1] makes this evident).

Smith suggests that a large bias was introduced by our procedure, in which respondents who indicated zero WTP for an inclusive good were assumed to have zero WTP for its constituents. We agree that the procedure could be tightened (see footnote 1 in the original paper). However, the possible bias was surely small in our study: the proportion of respondents indicating positive WTP for "improved availability of equipment and trained personnel" was 56, 58, and 61%, respectively, for the three groups in the study, in descending order of inclusiveness. Thus, the higher bound of the percentage of respondents to whom zero WTP could have been incorrectly imputed is 5%, clearly too little to change any conclusion of the study.

We disagree with Smith's insistence on the desirability of within-subject analyses. This suggestion misses the point of the study. We start from the premise that the current standards for CVM do not specify the appropriate level of embedding at which questions should be framed. Indeed, Smith never tells us what level of embedding would be appropriate in the study that we reported. We point out that this indeterminacy can have significant consequences, because different levels of embedding can yield systematically different estimates of value. Because a respondent in a CVM study is exposed only to a single framing, a between-subject design is most appropriate.

As we review Section 2 of Smith's article we are struck by what it does not say. Smith does not explain how the wording that we chose might have caused our results and provides no answer to the critical question: is there a wording of the requests presented to our groups 1 and 2 that would yield the same WTP as that which we observed in group 3? He commits himself to no predictions and offers no suggestions about empirical research that could lay the issue to rest. The logic of the criticism of our work is worth noting, because if pursued further it would tend to make CVM immune to empirical critique. In effect, the argument is that our study could have been done differently, that the outcome might then have been different, and that our results are therefore inconclusive. The problem with this mode of criticism is that it is too easy: it is true of every study that it could have been done differently. If you will only allow a completely conclusive study to change your mind, you will rarely have to do so.

Section 3 is mostly concerned with an attempt to show that our demonstrations of embedding are not necessarily incompatible with standard value theory. The broad lines of the argument are familiar and were anticipated in our article, where we discussed examples in which apparent embedding effects arise as a result of substitution and income effects. The critical task is not to show that substitution and income effects *could* produce the result, but to provide a plausible account, preferably supported by evidence, of their role in particular cases. This section does not do so, and it provides no answers to several critical questions. Does Smith endorse the hypothesis that the approximate constancy of WTP is caused by a particular pattern of income and substitution effects? How would this hypothesis be tested? Which level of embedding would be the appropriate one for a study of WTP for rescue equipment and personnel? Which specification of the timing of payments is appropriate for a contingent valuation? How should payment streams be described, if people do not spontaneously apply a consistent discounting model? Here again, the main claim is that our results are not fully conclusive, but there is no commitment to particular predictions, nor much evidence of interest in seeing the issues resolved by future research.

In summary, Smith's article does not lay to rest the question which our results brought into sharp focus and which had troubled many students of contingent valuation over the years: could embedding effects have been demonstrated in many —and perhaps all—of the "several hundred applications... completed to date" in which CVM was used to measure economic value? The results of a single study do not, of course, demonstrate that the valuation of every good will be susceptible to an embedding effect. However, together with earlier observations cited in our article, our results suggest that the likelihood of observing embedding effects is high, especially when the respondents have had no relevant market experience. In our view, the evidence appears sufficiently strong to impose on users of CVM the burden of demonstrating that their estimate of the value of a good is immune to manipulations of embedding. In the absence of such a demonstration of robustness, any assessment based on CVM should remain suspect. If embedding effects can be removed by improved survey techniques, the experiment that we have proposed will show this to be the case. The required tests are easily done, and their cost will usually be small relative to the stakes. The question of whether an estimate of the value of a particular public good is robust or arbitrary can be and should be settled by research, not by statements of faith.

Contrary to the implication of Smith's title, we did not propose a premature verdict against CVM. We do believe, however, that the premature verdict in its favor should be carefully reconsidered in a spirit of impartial scientific inquiry. We must voice our concern: it would be a truly appalling state of affairs if the wish to protect the rebuttable presumption of validity that CVM now enjoys in the courts were allowed to have a chilling effect on research.

REFERENCES

1. B. Fischhoff, Using surveys to value public goods: The contingent valuation method. *Public Opinion Quart.* 54, 286–289 (1990).

2. H. Fischhoff and L. Furby, Measuring values: A conceptual framework for interpreting transactions with special reference to contingent valuation of visibility, *J. Risk Uncertainty* 1, 147–184 (1988).

3. D. Kahneman and J. L. Knetsch, Valuing public goods: The purchase of moral satisfaction, *J. Environ. Econom. Management* 22, 57–70 (1992).

4. V. K. Smith, Arbitrary values, good causes, and premature verdicts, *J. Environ. Econom. Management* 22, 71–89 (1992).

ATTACHMENT N

ISSUES OF METHOD IN THE STUDY OF EMBEDDING: REPLY TO HARRISON

March 20, 1992

Daniel Kahneman
University of California, Berkeley

Jack Knetsch
Simon Fraser University

Harrison [3] accepts that the embedding effect is potentially "damaging to the applicability of the CVM" (p.248), but he insists that Kahneman and Knetsch [4, herafter KK] did not properly demonstrate embedding, supporting this claim by a reanalysis of the KK data. We find little merit in Harrison's critique: his methodological preferences are questionable, his reanalysis of our data is severely biased, and his alternative interpretation is devoid of any support. However, the effort to understand Harrison's position eventually led us to notice some significant differences among the questions that survey participants can be asked in studies of embedding. We focus our comments on this issue, rather than on an exhaustive list of Harrison's errors[1].

Three estimates of willingness to pay (WTP) for the same public good were obtained in the KK study, by presenting different sequences of questions to separate samples of randomly selected respondents in a telephone survey. Respondents in group 1 were asked their WTP "to improve environmental services"; the definition of this good explicitly included improved preparedness for disasters. Following Harrison's notation, the mean of these responses is labeled WTP(ABC1) (see Table I). Respondents who stated a positive WTP for environmental services were further asked an allocation question: "What part of the total amount that you just mentioned for all environmental services do you think should go specifically to improve preparedness for disasters?" [WTP(AB1)], and respondents who made a positive allocation to this good were asked to make a further allocation to one of its listed constituents, the improved availability "of equipment and trained personnel for rescue operations" [WTP(A1)]. Respondents in group 2 were asked their WTP for disaster preparedness [WTP(AB2)], and those who stated a positive WTP were asked for an allocation to equipment and trained personnel for rescue operations [WTP(A2)]. Finally, respondents in group 3 were simply asked to state their WTP for rescue equipment and trained personnel [WTP(A3)].

Our analysis of the data assumed that respondents who stated zero WTP for an inclusive good also assigned a zero allocation to constituents. Harrison objects to this procedure and offers an alternative analysis. To facilitate comparison, the main results of both analyses are presented in Table I.

INSERT TABLE 1 ABOUT HERE

In our view, the results of Table I illustrate two important phenomena. (i) The close similarity of WTP(ABC1) with WTP(AB2) and WTP (A3) is a case of the effect we have called *perfect embedding* [5]: similar estimates of WTP are observed for goods that vary greatly in inclusiveness. (ii) The large differences between the averages of WTP(A1), WTP(A2) and WTP(A3) illustrate a separate effect that we call *regular embedding*: different estimates of the same good are observed, depending on whether the good is evaluated on its own (A3) or as an allocation of WTP for a more inclusive good.

Harrison's view of our results and conclusions is different, and substantially incorrect. He organizes his discussion in terms of three claims that he attributes to KK. The *first claim* on Harrison's list is equivalent to our notion of perfect embedding, as defined above. There is no difference between KK and Harrison with respect to the empirical results: as indicated in Table I, estimates of the three WTP's (for ABC1, AB2 and A3) are identical in the two columns. The data clearly support the hypothesis of perfect embedding, but Harrison rejects it, with the odd argument that the result "has nothing to do with embedding" because KK asked "three *distinct groups of people* to evaluate three *distinct goods*" (italics in original). The objection to KK's conclusion is that it relies on a "missing, and rather unpalatable, assumption that all respondents are drawn at random from a given population". This was not an assumption: although we did not mention the technical details, the method of our experiment ensured random sampling from a common population. All the data summarized in Table I were obtained in one evening, and interviewers used a random device to select a questionnaire for each interview. Harrison appears to be generally uncomfortable with between-group comparisons, but the procedure employed by KK is routine in the biological and social sciences, where different treatments are most often applied to different subjects. In the context of contingent valuation, this standard experimental method is indispensable for the systematic investigation of alternative interviewing formats and other issues of technique. Tastes, incomes and other attributes of course vary among individuals, but it is safe to assume that substantial samples from the same population have similar distributions of these characteristics.

The *second claim* discussed by Harrison is simply a mistake on his part. Harrison attributes to KK the hypothesis that the allocations made to constituent goods after the original statement of WTP are very close to the original WTP. In support, he cites KK's statement that "the means of WTP across levels of embedding were also very close". Harrison evidently failed to note that this statement appears in a paragraph that documents perfect embedding (that is, the similarity of WTP values for ABC1, AB2 and A3 in Table 1). The meaning of the sentence is unambiguous in that context: it merely restates what Harrison calls the first claim. Harrison's interpretation is especially peculiar because the claim that he attributes to us is obviously falsified in the results we reported (see the KK column in Table I), and also incompatible with regular embedding, which was the main claim of the KK paper. We are forced to conclude that Harrison is confused about our central hypothesis. Fortunately, his misunderstanding appears to be idiosyncratic: we know of no one else who was similarly misled by the statement quoted above.

The *third claim* listed by Harrison is the effect that we now call regular embedding: average WTP for a particular good is lower when WTP is assessed by allocation of part of WTP for a more inclusive good than when willingness to pay for the particular good is elicited directly. Although Harrison's statistical analysis is severely biased, as we shall show below, its results confirm the hypothesis of regular embedding. Harrison nevertheless concludes with the statement, entirely unsupported by any evidence, that "one can either explain the KK results as due to sloppy instructions or as being perfectly consistent with the predictions of economic theory". We seriously doubt whether there is any result that could not be dismissed by assertions of this sort.

Although it did not eliminate regular embedding, Harrison's analysis of the KK data raises an issue of general interest. The discrepancies that show up in several rows of Table I reflect different treatments of respondents who stated zero WTP for an inclusive good. We imputed zero allocations to these respondents. Harrison professed to be shocked by this procedure, and he simply excluded these respondents from the analysis. To explain why our original choice was correct, we apply a conversational analysis that has become increasingly important in survey research in recent years [1,7]. The aim of this analysis is to identify the constraints that the wording of questions imposes on the range of possible answers.

The initial elicitation of WTP for an inclusive good can be followed by one of two possible types of questions about WTP for its constituents. Respondents may be asked an *allocation question*, as in the KK study. Alternatively, respondents may be asked an *independent question* about their maximum WTP for a constituent good, considered on its own.

Allocation questions are commonly used in CV studies [see, for example, 2, 8]. This question imposes a powerful constraint on responses because it makes no sense to allocate more than the amount previously mentioned. Indeed, *no* respondent in the KK study ever stated an allocation that exceeded the original contribution. The same conversational constraint applies to respondents who state zero WTP for the inclusive good: when there is nothing to be allocated, positive allocations make no sense. We did not ask respondents an allocation question after a zero response because of another conversational rule, which constrains the interviewer: it is a conversational violation to ask people a question for which there is only one possible answer.

Harrison's insistence that we should have asked a follow-up question even after a zero response could only be satisifed by asking an independent question, which does not raise these difficulties. However, this procedure would be improper, because subjects who initially stated zero WTP would be asked a different follow-up question than subjects who offered a positive contribution to the inclusive good. Harrison to the contrary notwithstanding, there appears to be no plausible alternative to the procedure used in KK, when allocation questions are used.

The general issue of whether embedding is better studied by allocation questions or by independent questions is beyond the scope of this reply. However, it is useful to note that our analysis implies contrasting patterns of responses to the two types of questions. As illustrated by the KK study, a pattern of initial contributions that approximates perfect embedding is sufficient (although not necessary) to ensure regular embedding when an allocation question is used, because respondents rarely allocate their entire initial contribution to a single constituent. When independent questions are used, however, regular embedding is demonstrated if the sum of WTP for separate constituents exceeds WTP for their sum, provided that the constituents are not substitutes. Furthermore, perfect embedding can be observed in the responses of single individuals. An illustrative demonstration is provided by the results of a small experiment conducted in May 1991.

Visitors to a science museum were invited to answer a brief questionnaire on environmental issues, for a payment of $2 and a matching donation to the museum. Participation was restricted to individuals who stated that they had filed an income tax return in the preceding year. The questions presented to the 103 respondents dealt with a special tax to achieve

> "... significant improvements in disaster preparedness in the Bay Area. Proposed changes include improvements in emergency services in hospitals, maintenance of large stocks of medical supplies, food, fuel and communication equipment, ensuring the availability of equipment and trained personnel for rescue operations, and preparations for cleanup of oil, toxic chemicals or radioactive materials"
>
> What is the most you would be willing to pay this year in higher taxes, to go into a special fund to improve disaster preparedness in the Bay Area? $______
>
> What is the most that you would be willing to pay this year in higher taxes, specifically to ensure the availability of equipment and trained personnel for rescue operations? $______"

The main difference between these questions and the questions used for group 2 of the KK study (see above) is that the second question of the new version can be interpreted as a request for an independent statement of WTP, not for an allocation of the amount indicated in answer to the first question. This seemingly minor alteration of wording had a striking effect. The independent question yielded almost perfect embedding in a within-subject design: mean WTP was $63 and $59, respectively for the first and the second questions. Of 80 respondents who stated positive WTP on the first question, 47 (59%) stated the same WTP for the second question, 10 (13%) stated a higher WTP for the second question than for the first, 21 (27%) stated a lower WTP, and 2 did not answer. Very different results had been obtained with the allocation question used in groups 1 and 2 of the KK study: 85 of 92 respondents who initially stated a positive WTP allocated a smaller amount to the good mentioned in the second question; 3 allocated the whole amount, and 4 refused to answer.

Even with the independent question used in the museum study, relatively few respondents stated a positive WTP for the more specific good after stating zero WTP for the inclusive one. Of the 23 respondents who indicated zero WTP on the first question, 4 (17%) indicated a positive WTP on the second question. This result implies that the estimates of WTP in the KK study would have been substantially the same even if an independent question had been used for this subset of respondents. A related observation is that the correlation between the responses to the two questions posed in the museum survey was very high (r = .96). This correlation induces a large systematic bias in Harrison's analysis of the KK data. Harrison's analysis compares mean WTP for rescue and equipment personnel in an intact group [WTP(A3)] and in groups from which he has excluded respondents who had zero WTP for a more inclusive good. The comparison is severely biased because the factor on which respondents are excluded is highly correlated with the dependent variable.

Is anything of substance left of Harrison's methodological argument or of his allegedly superior analysis of the KK data? We think not.

Finally, we turn to Harrison's scathing rejection of KK's hypothesis that WTP for public goods is (sometimes) best viewed as a purchase of moral satisfaction. Harrison is much fonder of an idea that he labels the Good Cause Dump hypothesis, to the effect that "agents can be viewed as having a basket of "good causes", and that when a CVM researcher comes along and asks them to value any constituent "good cause" in this basket they "dump" their whole valuation for the basket into the stated WTP" (p.249). The two hypotheses are obviously similar. One noteworthy difference between them is that the notion of moral satisfaction has empirical consequences: the KK study reported that independent ratings of the satisfaction of contributing to various public goods predicted the average WTP for these goods with fair accuracy. As it is stated by Harrison, it is not clear that the Good Cause Dump hypothesis has any empirical meaning at all. Without an independent assessment of the contents of the hypothesized basket of good causes, this hypothesis can explain any result, and predicts none. As in the rest of his critique, it would be a serious mistake to judge the quality of Harrison's arguments by the confidence with which they are asserted.

FOOTNOTES

* We thank Ralph d'Arge, Baruch Fischhoff, Robert Haveman, Richard Thaler, and especially Carol Nickerson for their comments. Carol Nickerson also carried out a thorough analysis of the data from our original study, and we borrow Table 1 from her. Karen Jacowitz and Paul Grant participated in the new embedding study described here.

1. Nickerson [6] documented Harrison's statistical and methodological errors.

REFERENCES

1. H.H. Clark and M.F. Schober, "Asking Questions and Influencing Answers." In J.M. Tanur, ed., Questions about Questions. New York: Russell Sage Foundation, 15-48, (1992).

2. W.H. Devousges, V.K. Smith and M.P. McGovney, "A Comparison of Alternative Approaches for Estimating Recreation and Related Benefits of Water Quality Improvements," Report to the U.S. Environmental Protection Agency, Washington, D.C. (1983).

3. G.W. Harrison, "Valuing Public Goods with the Contingent Valuation Method: A Critique of Kahneman and Knetch," Journal of Environmental Economics and Management 22, 248-257 (1992).

4. D. Kahneman and J.L. Knetsch, "Valuing Public Goods: The Purchase of Moral Satisfaction," Journal of Environmental Economics and Management 22, 7 (1992).

5. D. Kahneman and J.L. Knetsch, "Contingent Valuation and the Value of Public Goods: Reply, Journal of Environmental Economics and Management 22, 57-70 (1992).

6. C.A.E. Nickerson, "Valuing Public Goods: A Reply to Harrison's Critique of Kahneman and Knetsch," Department of Psychology, University of California-Berkeley, Working Paper (1991). **[A version of this paper is now in press in JEEM, under the title "Valuing Public Goods: A Comment on Harrison's Critique of Kahneman and Knetsch"]**

7. N. Schwarz, F. Strack and H.P. Mai, "Assimilation and Contrast in Part-Whole Question Sequences: A Conversational Logic Analysis, Public Opinion Quarterly 55, 3-23 (1991).

8. R.G. Walsh, J.B. Loomis and R.S. Gillman, "Valuing Option, Existence, and Bequest Demands for Wilderness," Land Economics 60, 14-29 (1984).

Table 1. A Comparison of Mean Valuations of Embedded Goods

	Kahneman and Knetsch		Harrison	
	N	Mean	N	Mean
ABC1	66	$135.91	66	$135.91
AB1	66	29.06	43	44.60
A1	64	14.42	41	22.51
AB2	78	151.60	78	151.60
A2	74	74.65	45	122.76
A3	74	122.64	74	122.64